Universidade de Aveiro Departamento de Matemática
2006

Ricardo Jorge Aparício Gonçalves Pereira

Polinómios Quaterniónicos e Sistemas Comportamentais

Quaternionic Polynomials and Behavioral Systems

Universidade de Aveiro Departamento de Matemática
2006

Ricardo Jorge Aparício Gonçalves Pereira

Polinómios Quaterniónicos e Sistemas Comportamentais

Quaternionic Polynomials and Behavioral Systems

dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Matemática, realizada sob a orientação científica da Doutora Maria Paula Macedo Rocha Malonek, Professora Catedrática do Departamento de Matemática da Universidade de Aveiro

Apoio financeiro da FCT, Bolsa de
Doutoramento SFRH / BD / 5213 / 2001

Para a Lara

o júri

presidente	Reitora da Universidade de Aveiro
vogais	Doutor José da Silva Lourenço Vitória Professor Catedrático Aposentado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra
	Doutor Fernando Abel da Conceição Silva Professor Catedrático da Faculdade de Ciências da Universidade de Lisboa
	Doutor Jan Willems Professor Catedrático do Department of Electrical Engineering, Katholieke Universiteit Leuven, Bélgica
	Doutora Maria Paula Macedo Rocha Malonek Professora Catedrática da Universidade de Aveiro (Orientadora)
	Doutor João Carlos David Vieira Professor Associado Aposentado da Universidade de Aveiro
	Doutor Paolo Vettori Professor Auxiliar Convidado da Universidade de Aveiro

agradecimentos

À minha orientadora Doutora Paula Rocha Malonek que, desde o primeiro dia, me incutiu o espírito da investigação matemática. Nunca esquecerei a sua constante disponibilidade, apoio, incentivo e amizade. A ela o meu sincero agradecimento por tudo.

Ao Doutor Paolo Vettori pela colaboração útil e pelo partilhar de ideias das quais surgiram alguns resultados deste trabalho.

Ao Departamento de Matemática da Universidade de Aveiro por me ter concedido excelentes condições de trabalho durante estes últimos quatro anos. E a todos os meus colegas do departamento, incluindo funcionários não docentes, pelo incentivo e ambiente de trabalho proporcionado.

À Fundação para a Ciência e Tecnologia pelo apoio financeiro que me prestou.

Aos meus pais e à minha irmã por todo o carinho ao longo da minha vida. Sem o vosso apoio seria impossível chegar aqui.

À Ricarda que, especialmente neste último ano, esteve sempre presente, foi compreensiva e me soube transmitir toda a confiança necessária para ultrapassar as fases menos boas desta longa caminhada.

palavras-chave Quaterniões, Polinómios, Matrizes, Sistemas Comportamentais

resumo

Nesta tese consideramos sistemas comportamentais cujas trajectórias são as soluções de equações às diferenças ou diferenciais quaterniónicas. Como acontece no caso real ou complexo, os polinómios quaterniónicos e as matrizes quaterniónicas desempenham um papel muito importante na caracterização das propriedades dinâmicas de tais comportamentos. Portanto, uma grande parte desta tese é dedicada ao estudo destes objectos.

Depois de darmos algumas noções sobre quaterniões e matrizes quaterniónicas, introduzimos os polinómios quaterniónicos. A relação não trivial entre zeros e divisores destes polinómios é analisada. Mais ainda, algumas ferramentas algébricas tais como o determinante polinomial e a forma de Smith quaterniónica de matrizes polinomiais quaterniónicas são apresentadas e caracterizadas. Estas ferramentas são essenciais na caracterização da controlabilidade e da estabilidade de comportamentos quaterniónicos. Também introduzimos a forma de Smith-McMillan de matrizes racionais quaterniónicas, que é usada no estudo da estabilidade BIBO de um comportamento quaterniónico.

keywords Quaternions, Polynomials, Matrices, Behavioral Systems

abstract

In this thesis we considerer behavioral systems whose trajectories are given as solutions of quaternionic difference or differential equations. As happens in the real or complex case, it turns out that quaternionic polynomials and polynomial matrices play an important role in the characterization of the dynamical properties of such behaviors. Therefore, great part of this thesis is devoted to the study of those objects.

After giving some notions on quaternions and quaternionic matrices we introduce the quaternionic polynomials. The non-trivial relation between zeros and divisors of these polynomials is analyzed. Moreover, some algebraic tools such as the polynomial determinant and the quaternionic Smith form of quaternionic polynomial matrices are presented and characterized. These tools are essential in the characterization of the controllability and stability of quaternionic behaviors. We also introduce the quaternionic Smith-McMillan form of quaternionic rational matrices, which is used in the study the BIBO-stability of a quaternionic behavior.

Contents

Introduction

Quaternionic numbers, or simply quaternions, were first introduced by the Irish Mathematician Sir William R. Hamilton in [19] as a generalization of the complex numbers that he had previously studied in [18]. Basically, quaternions differ from complex numbers since they are described in terms of three, instead of one, noncommuting imaginary units.

As expected, Hamilton's first attempt to generalize his complex theory was to consider numbers corresponding to triplets instead of to pairs (as is the case of the complex numbers). However it was not possible to properly define the product of such numbers in order to guarantee a desired isomorphism between $\{(x, y, 0) : x, y \in \mathbb{R}\}$ and $\{(x, y) : x, y \in \mathbb{R}\} \simeq \mathbb{C}$. This major drawback could only be solved by an increase of dimension, giving rise to the quaternions. The set of all quaternions is usually denoted by $\mathbb{H}$.

These numbers may be favorably used to describe phenomena occurring in areas such as electromagnetism and quantum physics [50] by means of a compact notation that leads to a higher efficiency in computational terms [21].

In particular, quaternions are a powerful tool in the description of rotations. Indeed, by identifying $\mathbb{R}^3$ with a subset of $\mathbb{H}$, the rotation of a vector $v \in \mathbb{R}^3$ by a given angle about a certain direction can be expressed as qvq^{-1}, where q is a suitable quaternion (see, e.g., [26]).

It is not uncommon to be faced with situations, especially in robotics, where the rotation of a rigid body depends on time, and this dynamics is advantageously written in terms of quaternionic differential or difference equations. The effort to control the rotation dynamics motivates the study of these equations from a system theoretic point of view (see, for instance, [12]).

In the context of quantum mechanics, a possible quaternionic formulation of the Schrödinger equation has been proposed since the sixties when experiments to check the existence of quaternionic potentials where performed (see, for instance, [31]). This theory leads to differential equations with quaternionic coefficients [9].

Using quaternionic notation it is also possible to find an elegant solution of the differential equation which describes the orbits of the planets, i.e., to solve the "Kepler problem" (see [51]). Quaternions, compared to vectors in $\mathbb{R}^3$, have an extra degree of freedom which may be exploited to simplify the equations. Indeed, by choosing conveniently the free parameter, the problem is reduced to the solution of the simple quaternionic differential equation $\ddot{q} + q = 0$. The solution of this equation is $q = e^{\mathbf{i}t}\alpha + e^{-\mathbf{i}t}\beta$, as in the commutative case, but where now α and β are constant quaternions.

Our work is motivated by the study of dynamical systems described by quaternionic differential or difference equations. Similarly to what happens in the real and complex cases, quaternionic polynomials an polynomial matrices are an important tool in the analysis of such systems and deserve therefore great attention in this thesis.

Although at this stage we avoid entering into details, we would like to note that there is no unique way of defining quaternionic polynomials (see, e.g., [44]). Actually, letting $a, b \in \mathbb{H}$ and s be the indeterminate, the monomials

$$asb, \qquad abs \qquad \text{and} \qquad sab$$

are different due to the noncommutativity of the quaternions. In this thesis we shall define polynomials that formally resemble real or complex polynomials since the indeterminate behaves as if it *commuted* with the coefficients. The reason of this choice lies in the deep connection between such polynomials and linear difference and differential equations with quaternionic coefficients, where the coefficients commute with the elementary difference and differential operators.

In spite of resemblance, these polynomials are noncommuting and differ substantially from the complex ones as regards zeros and factors. Indeed, some of their properties which were already studied in [16, 27, 35] may even seem surprising at first sight.

The first three chapters give an overview of the main relevant results on quaternions and quaternionic polynomials and polynomial matrices. Besides presenting well-known material in this area, we introduce new concepts and develop some new results that are relevant for our purposes. More concretely, we propose a definition of determinant for quaternionic polynomial matrices, give an alternative characterization of Smith-forms, and introduce Smith-McMillan forms. After having set the algebraic tools, we dedicate the last two chapters to the study of quaternionic dynamical systems within the behavioral framework.

The behavioral approach to dynamical systems, introduced by J. C. Willems [53, 54] in the eighties, considers as the main object of study in a system the set of all the trajectories which are compatible with its laws, known as the system behavior. Whereas the classical approaches start by dividing the trajectories into input, output and/or state space variables, according to some predefined mathematical model (for instance, the input-output or the state space model), the point of view of the behavioral approach is rather innovative. One looks at the set of trajectories without imposing any structure, i.e., without speaking, at an early stage, of inputs and outputs, of causes and effects. This point of view does not only unify the previous approaches, fitting them within an elegant theory, but it also permits to study a larger class of dynamical systems, including situations where it is not possible or desirable to make any distinction between input and output variables.

Systems with quaternionic signals were already investigated in the classic state space approach [20]. Here we aim at laying the foundations of the theory of quaternionic systems in the behavioral approach. Although every quaternionic system can be regarded as a complex or real system of higher dimension with special structure, keeping at the quaternionic level (i.e., viewing it as a system over $\mathbb{H}$) allows higher efficiency in computational terms.

In order to make this thesis as far as possible self-contained, we decided to include some proofs of non-original results. If nothing is mentioned, the reference for the proof is the same as the one of the corresponding result. All the other results and their proofs are original and are most of them contained in our papers [40, 41, 42].

We conclude the introduction with a brief outline of the contents of each chapter of this thesis.

Chapter 1 - Quaternions and quaternionic matrices

In this chapter, after giving some basic notions about quaternions, several alternative representations of quaternions are derived and the relation of similarity is defined and characterized. Then quaternionic matrices as well as complex adjoint matrices are introduced and the issue of determinants is addressed. Namely, two determinants of quaternionic matrices are introduced, the Dieudonné determinant and the Study determinant. After analyzing the eigenvalue problem, the chapter is concluded with an example of application of the quaternions.

Chapter 2 - Quaternionic polynomials

After introducing the quaternionic polynomials, some basic notions, such as zeros and divisors, are given and characterized. It turns out that the relation between zeros and divisors is not as simple as in the commutative case, which leads to some surprising results. Moreover, a new definition of total divisor is introduced, which proves to be equivalent to other existing definitions and characterizations [23]. Finally, the similarity of quaternionic polynomials is studied.

Chapter 3 - Quaternionic polynomial and rational matrices

This chapter starts with some definitions and preliminary results on quaternionic polynomials and rational matrices, most of which are simply an extension of the ones for real or complex matrices. Then, following as close as possible the approach of Dieudonné to quaternionic (constant) matrices a new definition of determinant for quaternionic polynomial matrices is given. Complex adjoint matrices are defined and shown to share many algebraic properties with the corresponding quaternionic polynomial matrices.

In Section 3.4 we introduce the quaternionic Smith form of quaternionic polynomial matrices, which unlike the commutative case is not unique. We also characterize the complex Smith form of complex adjoint matrices and give its relation with the quaternionic Smith form. Finally, in Section 3.5, we introduce the quaternionic Smith-McMillan form of a quaternionic rational matrix and the results of the previous section are extended to rational matrices.

Chapter 4 - Quaternionic behavioral systems

The behavioral approach introduced by Willems [53, 54] is extended here to quaternionic systems. We start by studying behaviors that can be described as solution sets of quaternionic matrix difference or differential equations, i.e., those which are the kernel of some suitable matrix difference or differential operator. Difference equations arise either directly or from the digital implementation of problems described by differential equations. We show that as in the real and in the complex case, two matrices represent the same behavior if and only if each one is a left multiple of the other. Then we also analyze other representations of a system, such as image and input-output representations. Finally we give a complete and explicit characterization of all solutions of quaternionic matrix difference and differential equations.

Chapter 5 - Dynamical properties of quaternionic behaviors

In this chapter it is shown how basic but fundamental dynamical properties of a quaternionic behavior such as controllability and stability can be characterized in terms of its kernel representations. It turns out that the previously introduced algebraic tools, such as the quaternionic Smith and Smith-McMillan forms and the determinant of quaternionic polynomial matrices, play an important role in these characterizations.

Chapter 1

Quaternions and quaternionic matrices

In this chapter, after giving some basic notions about quaternions, several alternative representations of quaternions are derived. Since the product of quaternions is non-commutative, some well-known concepts can not be defined as in the real or complex case. For instance, the usual notion of determinant of real matrices can not be extended to quaternionic matrices. It turns out that different but closely related definitions of determinant of quaternionic matrices can be given. Moreover, it is necessary to make the distinction between left and right eigenvalues and many results concerning right eigenvalues, which are the most useful for our purposes, are stated. Finally it is shown how the rotation of a rigid body can be described in a quaternionic framework.

1.1 Quaternions

Quaternions were first introduced in 1843 by the Irish Mathematician Sir William R. Hamilton, motivated by giving a generalization of complex numbers [19]. A quaternion is a number of the form

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \ : \ a, b, c, d \in \mathbb{R}, \tag{1.1}$$

where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, the *imaginary units*, commute with real numbers and satisfy

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

This implies that

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \qquad \mathbf{ki} = -\mathbf{ik} = \mathbf{j},$$

and therefore the following multiplication table is obtained

	1	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
1	1	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{i}$	$\mathbf{i}$	-1	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$\mathbf{j}$	$-\mathbf{k}$	-1	$\mathbf{i}$
$\mathbf{k}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{i}$	-1

The set of all quaternions is denoted by $\mathbb{H}$. Note that $\mathbb{H}$ is a skew-field, also called associative division algebra [29], since it is noncommutative, as can be seen in the previous table. Moreover, Frobenius proved in 1878 that $\mathbb{H}$ is the only finite-dimensional associative division algebra over $\mathbb{R}$ of higher dimension than two [14].

The *conjugate* of $\eta = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$ is the quaternion

$$\overline{\eta} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}.$$

The mapping $\eta \mapsto \overline{\eta}$ is called *conjugation* and has the following properties:

$$\begin{array}{llll} (i) & \overline{x\eta + y\nu} = x\overline{\eta} + y\overline{\nu} & (ii) & \overline{\overline{\eta}} = \eta \\ (iii) & \eta\overline{\eta} = \overline{\eta}\eta & (iv) & \overline{\eta\nu} = \overline{\nu}\,\overline{\eta}, \end{array}$$

for all $\eta, \nu \in \mathbb{H}$ and $x, y \in \mathbb{R}$.

If we define the *real* and *imaginary parts* of η as $\operatorname{Re}\eta = a$ and $\operatorname{Im}\eta = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, respectively, many formulæ which are analogous to the complex case can be written. For instance,

$$\overline{\eta} = \operatorname{Re}\eta - \operatorname{Im}\eta, \quad \operatorname{Re}\eta = \frac{1}{2}(\eta + \overline{\eta}) \quad \text{and} \quad \operatorname{Im}\eta = \frac{1}{2}(\eta - \overline{\eta}).$$

If $\operatorname{Re}\eta = 0$, η is called a *pure quaternion*.

The *norm* of the quaternion $\eta = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ is defined as

$$|\eta| = \sqrt{\overline{\eta}\eta} = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

This is indeed a norm in $\mathbb{H}$ since, for all $\eta, \nu \in \mathbb{H}$, it satisfies

$$(i)\ |\eta| = 0 \Leftrightarrow \eta = 0 \qquad (ii)\ |\eta + \nu| \le |\eta| + |\nu| \qquad (iii)\ |\eta\nu| = |\eta||\nu|.$$

Note that $|\eta|^2 = \operatorname{Re}^2\eta + |\operatorname{Im}\eta|^2$. A quaternion η is said to be *unitary* if $|\eta| = 1$.

Each nonzero $\alpha \in \mathbb{H}$ has an *inverse* given by $\alpha^{-1} = \frac{\overline{\alpha}}{|\alpha|^2}$. The inverse of the product is given by $(\alpha\beta)^{-1} = \beta^{-1}\alpha^{-1}$, for all nonzero $\alpha, \beta \in \mathbb{H}$.

A concept which will play an important role throughout this thesis is the similarity relation. This relation is defined and characterized below.

Definition 1.1.1. [55] Two quaternions η and ν are said to be *similar*, $\eta \sim \nu$, if there exists a nonzero $\alpha \in \mathbb{H}$ such that $\eta = \alpha\nu\alpha^{-1}$.

Similarity is an equivalence relation and we denote by $[\nu]$ the equivalence class containing ν.

In general, it is not easy to check by the definition whether two quaternions are similar or not. However, an easy characterization of similarity can be given.

Proposition 1.1.2. [4, 55] *Two quaternions η and ν are similar if and only if*

$$\operatorname{Re}\eta = \operatorname{Re}\nu \quad \textit{and} \quad |\eta| = |\nu|.$$

Therefore, for instance, all the imaginary units belong to the same equivalence class, i.e., $\mathbf{i} \sim \mathbf{j} \sim \mathbf{k}$. Furthermore, for all $\eta \in \mathbb{H}$,

$$\eta \sim \overline{\eta}.$$

As a consequence of the characterization of similarity, it turns out that every equivalence class in $\mathbb{H}/\sim$ contains a complex representative, as stated in the following proposition. First note that the set $\{a + b\mathbf{i} : a, b \in \mathbb{R}\} \subseteq \mathbb{H}$ is isomorphic to the complex field, and therefore we will implicity assume throughout this thesis that

$$\mathbb{C} = \{a + b\mathbf{i} : a, b \in \mathbb{R}\}.$$

Proposition 1.1.3. [55] *For all $\eta \in \mathbb{H}$, $[\eta] \cap \mathbb{C} \neq \emptyset$.*

Proof. Let $\eta = a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ be an arbitrary quaternion and $z = a+\sqrt{b^2 + c^2 + d^2}\mathbf{i} \in \mathbb{C}$. It is immediately seen that $\operatorname{Re}\eta = \operatorname{Re}z$ and $|\eta| = |z|$ and therefore, by Proposition 1.1.2, $\eta \sim z$. $\square$

Note that, given a quaternion η, there are at most two complex numbers in the equivalence class $[\eta]$, namely z and $\overline{z}$. Moreover, $[\eta] = \{\eta\}$ if and only if $\eta \in \mathbb{R}$. On the other hand, if $\eta \in \mathbb{H} \setminus \mathbb{R}$ then the equivalence class of η has infinite elements.

Example 1.1.4. Consider the quaternion $\eta = 1 - 2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$. The complex $z = 1 + 3\mathbf{i}$ and its conjugate $\overline{z} = 1 - 3\mathbf{i}$ are similar to η, since $\operatorname{Re}z = \operatorname{Re}\overline{z} = \operatorname{Re}\eta = 1$ and $|z| = |\overline{z}| = |\eta| = \sqrt{10}$. Furthermore,

$$[\eta] = \{\nu = 1 + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} : \ b, c, d \in \mathbb{R} \ \text{ and } \ b^2 + c^2 + d^2 = 9\}.$$

$\square$

Hamilton's original definition based on the imaginary units $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, presented in expression (1.1) is known as the *algebraic representation* of a quaternion. There are however other forms of representing quaternions.

For instance, identifying the basis elements of $\mathbb{R}^4$, $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$ with 1, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, respectively, the quaternion $\eta = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ can be viewed as the element (a, b, c, d) of the vector space $\mathbb{R}^4$.

An alternative representation, which will be the basis for the matrix representations presented further is the following. Since $\mathbf{ij} = \mathbf{k}$, given a quaternion $\eta = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ it is possible to write

$$\eta = z + w\mathbf{j},$$

where $z = a + b\mathbf{i} \in \mathbb{C}$ and $w = c + d\mathbf{i} \in \mathbb{C}$. This is called the *complex representation* of the quaternion η.

Using this notation, the conjugate of the quaternion η can be written as

$$\overline{\eta} = \overline{z} - w\mathbf{j}.$$

Note that the imaginary unit $\mathbf{i}$ commutes with the complex numbers while the imaginary units $\mathbf{j}$ and $\mathbf{k}$ do not. In fact, for all $z \in \mathbb{C}$, the following relations hold

$$\mathbf{i}z = z\mathbf{i}, \qquad \mathbf{j}z = \overline{z}\mathbf{j}, \qquad \mathbf{k}z = \overline{z}\mathbf{k}. \tag{1.2}$$

Similar to what happens in the complex case it is also possible to derive a trigonometric representation for quaternions. The description of rotations in a quaternionic framework, that will be presented on section 1.3, makes use of such representation.

Let then η be a quaternion with nonzero imaginary part. Clearly

$$\begin{aligned} \eta = \operatorname{Re}\eta + \operatorname{Im}\eta &= |\eta| \left(\frac{\operatorname{Re}\eta}{|\eta|} + \frac{\operatorname{Im}\eta}{|\eta|} \right) \\ &= |\eta| \left(\frac{\operatorname{Re}\eta}{|\eta|} + \frac{\operatorname{Im}\eta}{|\operatorname{Im}\eta|} \frac{|\operatorname{Im}\eta|}{|\eta|} \right). \end{aligned} \tag{1.3}$$

Since

$$\left(\frac{\operatorname{Re}\eta}{|\eta|} \right)^2 + \left(\frac{|\operatorname{Im}\eta|}{|\eta|} \right)^2 = 1,$$

it is natural to identify the quantities $(\operatorname{Re}\eta/|\eta|)$ and $(|\operatorname{Im}\eta|/|\eta|)$ with the cosine and sine of an angle θ, respectively. Moreover, the quaternion $u = \frac{\operatorname{Im}\eta}{|\operatorname{Im}\eta|}$ is unitary. Hence, equation (1.3) can be written as

$$\eta = |\eta|(\cos\theta + u\sin\theta), \qquad \theta \in]0,\pi[. \tag{1.4}$$

This is called the *trigonometric representation* of a quaternion.

Note that, if $\operatorname{Im}\eta = 0$, then the trigonometric representation of η is simply $\eta = |\eta|\cos 0$ or $\eta = |\eta|\cos\pi$, depending on the sign of η.

Remark 1.1.5. As is well-known, in the complex case, the trigonometric representation of $z \in \mathbb{C}$ is given by $z = |z|(\cos\theta + \mathbf{i}\sin\theta)$, $\theta \in [0, 2\pi[$, i.e., the angle θ and the norm $|z|$ uniquely determine z. By equation (1.4), it turns out that, besides the angle and the norm, also the unit quaternion u is necessary in the trigonometric representation of a quaternion. □

It is clear that the conjugate of the quaternion η defined as in (1.4) has the trigonometric representation

$$\eta = |\eta|(\cos\theta - u\sin\theta), \qquad \theta \in]0,\pi[.$$

Example 1.1.6. The trigonometric representation of the quaternion $\eta = 2+\mathbf{i}+\frac{\sqrt{2}}{2}\mathbf{j}+3\mathbf{k}$ is

$$\eta = 4\left(\cos\frac{\pi}{3} + u\sin\frac{\pi}{3}\right), \quad \text{with } u = \frac{\sqrt{3}}{6}\mathbf{i} + \frac{\sqrt{6}}{12}\mathbf{j} + \frac{\sqrt{3}}{2}\mathbf{k}.$$

□

We have seen so far one vector and three scalar (algebraic, complex and trigonometric) representations of a quaternion. Next, a complex and a real matrix representation are given. The first one [24, 55] associates the quaternion $\eta = z + w\mathbf{j}$, where $z = a + b\mathbf{i} \in \mathbb{C}$ and $w = c + d\mathbf{i} \in \mathbb{C}$, with the complex matrix

$$\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}. \tag{1.5}$$

This representation is called the *complex matrix representation* of a quaternion. Note that the complex matrix representation of the conjugate of η, $\overline{\eta} = \overline{z} - w\mathbf{j}$, is

$$\begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix}, \tag{1.6}$$

i.e., it is the transpose conjugate of the matrix given in (1.5).

Since the complex numbers are isomorphic to the 2×2 real matrices of the form

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \tag{1.7}$$

it is possible to associate the quaternion η with a real matrix. Indeed, replacing each complex number in (1.5) by its correspondent real matrix as in (1.7) we get the *real matrix representation* of η

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}.$$

It is not difficult to check that the conjugation of a quaternion corresponds simply to a transposition of its real matrix representation.

1.2 Quaternionic matrices

In this section we introduce matrices with quaternionic entries and analyze to which extent properties of matrices over fields can be generalized for this matrices.

Let $\mathbb{H}^{m\times n}$ denote the collection of all $m \times n$ matrices with quaternionic entries. To simplify the notation, if $n = 1$, then we identify $\mathbb{H}^{m\times 1} \simeq \mathbb{H}^m$.

The nomenclature adopted in the quaternionic case is similar to the one used for complex matrices.

Given $A = (a_{st}) \in \mathbb{H}^{m\times n}$ its *conjugate* is defined as $\overline{A} = (\overline{a}_{st})$, its *transpose* as $A^T = (a_{ts}) \in \mathbb{H}^{n\times m}$, and its *conjugate transpose* as $A^* = \overline{A}^T \in \mathbb{H}^{n\times m}$.

A square matrix $A \in \mathbb{H}^{n\times n}$ is *invertible* if there exists a matrix $B \in \mathbb{H}^{n\times n}$ such that $AB = BA = I_n$, where I_n is the $n \times n$ identity matrix. If A is not invertible, then A

is said to be *singular*.

The noncommutativity of quaternions suggests that some usual properties of matrices over a field do not hold in the quaternionic case. For instance, given two quaternionic matrices of suitable dimensions A and B it may happen that $(AB)^T \neq B^T A^T$. Nevertheless, the following properties hold.

Proposition 1.2.1. [55] *Let $A \in \mathbb{H}^{m\times n}$ and $B \in \mathbb{H}^{n\times p}$. Then*

1. $\left(\overline{A}\right)^T = \overline{A^T}$;
2. $(AB)^* = B^* A^*$;
3. $(AB)^{-1} = B^{-1}A^{-1}$ *if A and B are invertible;*
4. $(A^*)^{-1} = (A^{-1})^*$ *if A is invertible.*

According to [1], we shall use the following notation which will be necessary in the sequel (Section 1.2.2).

Notation 1.2.2. Denote by P_{lm} the matrix that is obtained from the identity by interchanging the l^{th} and m^{th} rows. Denote by $B_{lm}(\alpha)$, where $\alpha \in \mathbb{H}$, the matrix that is obtained from the identity by adding the m^{th} row multiplied by α to the l^{th} row. Finally denote by $SL(n,\mathbb{H})$ the set of all $n \times n$ matrices that can be decomposed as a finite product of matrices of the types P_{lm} and $B_{lm}(\alpha)$, $\alpha \in \mathbb{H}$.

1.2.1 Complex adjoint matrix

Similar to what happens with quaternions, each quaternionic matrix $A \in \mathbb{H}^{m\times n}$ can be uniquely written as a sum $A = A_1 + A_2\mathbf{j}$, where A_1, $A_2 \in \mathbb{C}^{m\times n}$. Therefore, we can define an injective $\mathbb{R}$-linear map: $\mathbb{H}^{m\times n} \to \mathbb{C}^{2m\times 2n}$ such that

$$A \mapsto A^c = \begin{bmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{bmatrix}. \tag{1.8}$$

The matrix A^c is called the *complex adjoint matrix* of A. In general, any complex matrix with the structure shown in (1.8) is said to be a *complex adjoint matrix.*

The representation via complex adjoint matrices, firstly introduced by Lee in 1949 [30], is one of the most efficient ways to study quaternionic matrices. Indeed, many of the results concerning quaternionic matrices that will be given in this thesis are proved applying complex adjoint matrices. Moreover, in Chapter 3 similar procedures are applied for results on quaternionic polynomial matrices

We may as well define a bijective $\mathbb{R}$-linear map: $\mathbb{H}^{m\times n} \to \mathbb{C}^{2m\times n}$ such that

$$A \mapsto A^{\mathbb{C}} = \begin{bmatrix} A_1 \\ -\overline{A}_2 \end{bmatrix}, \tag{1.9}$$

which, in particular, maps column vectors into column vectors. This is an isometry of the vector spaces $\mathbb{H}^m$ and $\mathbb{C}^{2m}$, i.e., $||v|| = ||v^{\mathbb{C}}||$, $\forall v \in \mathbb{H}^m$, where $||.||$ denotes the usual norm of a vector.

The following properties hold for complex adjoint matrices.

Proposition 1.2.3. [55] *Let A and B be quaternionic matrices of suitable dimensions and, in case, invertible. Then*

1. $(I_n)^c = I_{2n}$;
2. $(A^{-1})^c = (A^c)^{-1}$;
3. $(AB)^c = A^cB^c$;
4. $(AB)^{\mathbb{C}} = A^cB^{\mathbb{C}}$.

Remark 1.2.4. In an analogous way, given $A \in \mathbb{H}^{m\times n}$ we may write

$$A = A_1 + A_2\mathbf{j} = (A_{11} + A_{12}\mathbf{i}) + (A_{21} + A_{22}\mathbf{i})\mathbf{j}$$

and define an injective $\mathbb{R}$-linear map: $\mathbb{H}^{m\times n} \to \mathbb{R}^{4m\times 4n}$ such that

$$A \mapsto A^r = \left[\begin{array}{cc|cc} A_{11} & A_{12} & A_{21} & A_{22} \\ -A_{12} & A_{11} & -A_{22} & A_{21} \\ \hline -A_{21} & A_{22} & A_{11} & -A_{12} \\ -A_{22} & -A_{21} & A_{12} & A_{11} \end{array}\right].$$

The real matrix A^r is known as the *real adjoint matrix* of A. □

1.2.2 Determinants

Let $A \in \mathbb{R}^{n\times n}$ and denote by A_l, $l = 1, \dots, n$, the columns of A, i.e., $A = \begin{bmatrix} A_1 | \cdots | A_n \end{bmatrix}$. It is well know that the determinant of a real matrix satisfies, for instance, the following properties [22]

i) $\det\left(\begin{bmatrix} A_1 | \cdots | \alpha A_l | \cdots | A_n \end{bmatrix}\right) = \alpha \det\left(\begin{bmatrix} A_1 | \cdots | A_l | \cdots | A_n \end{bmatrix}\right), \quad \alpha \in \mathbb{R};$

ii) $\det I = 1$, where I is the identity matrix.

Due to the noncommutativity of $\mathbb{H}$, it is not possible to extend the usual definition of determinant to quaternionic matrices. For example, let

$$A = \begin{bmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{j} \end{bmatrix},$$

and suppose that the previous properties *i)* and *ii)* hold for quaternionic matrices. Then

$$\det A = \det \begin{bmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{j} \end{bmatrix} = \mathbf{i} \det \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{j} \end{bmatrix} = \mathbf{ij} \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{ij} = \mathbf{k}$$

whereas, on the other hand

$$\det A = \det \begin{bmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{j} \end{bmatrix} = \mathbf{j} \det \begin{bmatrix} \mathbf{i} & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{ji} \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{ji} = -\mathbf{k},$$

leading to an absurd.

The first mathematician who tried to define the determinant of a quaternionic matrix was Arthur Cayley in 1845 [5], but his definition was not satisfactory. Only in the twentieth century new developments in this topic were achieved and some different definitions such as the determinants of Dieudonné [11] and Study [47] were given.

These concepts are in accordance with the following definition of determinant for quaternionic matrices, that can be regarded as a generalization of the notion of determinant for the real and complex cases.

Definition 1.2.5. [1, 7] A function $d : \mathbb{H}^{n\times n} \to \mathbb{H}$ is said to be a *determinant* if it satisfies the following axioms:

(i) $d(A) = 0$ if and only if A is singular.

(ii) $d(AB) = d(A)d(B)$ for all $A, B \in \mathbb{H}^{n\times n}$.

(iii) If $A' = B_{lm}(\alpha)A$, $\alpha \in \mathbb{H}$, then $d(A') = d(A)$.

We next present the definition of determinant given by Dieudonné. For that purpose, we first give an auxiliary Lemma.

Lemma 1.2.6. [1, 11] *Let $A \in \mathbb{H}^{n\times n}$ be invertible. Then there exists a matrix $U \in SL(n, \mathbb{H})$ such that*

$$UA = \mathrm{diag}(1, \dots, 1, \alpha) \in \mathbb{H}^{n\times n}, \ \ \alpha \in \mathbb{H}.$$

Definition 1.2.7. [1, 11] Let $A \in \mathbb{H}^{n\times n}$; the *Dieudonné determinant* of A, denoted by $\mathrm{Ddet}(A)$, is defined as follows.

- If A is singular, then $\mathrm{Ddet}(A) = 0$.
- Otherwise, let $U \in SL(n, \mathbb{H})$ be such that

$$UA = \mathrm{diag}(1, \dots, 1, \alpha) \in \mathbb{H}^{n\times n}, \ \ \alpha \in \mathbb{H}.$$

  Then $\mathrm{Ddet}(A) = |\alpha|$.

Remark 1.2.8. Note that, as expected, the Dieudonné determinant is not an extension of the determinant of real matrices, i.e., given a real matrix $A \in \mathbb{R}^{n\times n}$ then, in general, $\mathrm{Ddet}(A) \neq \det(A)$. For instance, consider the simple scalar case $A = -1$. Then $\det(A) = -1$ but $\mathrm{Ddet}(A) = |-1| = 1$. □

Another definition of determinant of quaternionic matrices was proposed by Study in [47]. This determinant is defined in terms of complex adjoint matrices as follows.

Definition 1.2.9. [1, 47] Let $A \in \mathbb{H}^{n\times n}$, the *Study determinant* of A (also referred to as *q-determinant* in [55]) is defined as $\mathrm{Sdet}(A) = \det(A^c)$, where A^c is the complex adjoint matrix of A.

Based on this definition of determinant the well know Cayley-Hamilton Theorem was extended to the quaternionic case.

Proposition 1.2.10. [55] *Let $A \in \mathbb{H}^{n\times n}$ and define the characteristic polynomial of A as $F_A(\lambda) = \det(\lambda I_{2n} - A^c) \in \mathbb{C}[\lambda]$. Then $F_A(A) = 0$.*

Since each matrix $A \in \mathbb{H}^{n\times n}$ is also associated to a real adjoint matrix A^r, it is natural to consider as well the determinant $\det(A^r)$.

The aforementioned determinants are closely related as shown in the following proposition.

Proposition 1.2.11. [1, 2] *Let $A \in \mathbb{H}^{n\times n}$. Then*

$$\det(A^r) = \left[\mathrm{Sdet}(A)\right]^2 = \left[\mathrm{Ddet}(A)\right]^4.$$

Example 1.2.12. Let

$$A = \begin{bmatrix} -2\mathbf{k} & 2 \\ 2 - 3\mathbf{i} + 2\mathbf{j} & 2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k} \end{bmatrix} \in \mathbb{H}^{2\times 2}.$$

The matrix

$$U = \begin{bmatrix} 1 & 0 \\ \mathbf{i} - 2\mathbf{j} & 1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{i} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\mathbf{j} - \mathbf{k} & 1 \end{bmatrix} = \begin{bmatrix} 1 + \mathbf{j} - \mathbf{k} & \mathbf{i} \\ 2 + 3\mathbf{i} - 2\mathbf{j} & 2\mathbf{k} \end{bmatrix} \in SL(2, \mathbb{H})$$

is such that $UA = \begin{bmatrix} 1 & 0 \\ 0 & 2\mathbf{i} \end{bmatrix}$ and hence $\mathrm{Ddet}(A) = |2\mathbf{i}| = 2$.

Moreover, since

$$A^c = \begin{bmatrix} 0 & 2 & -2\mathbf{i} & 0 \\ 2-3\mathbf{i} & 2\mathbf{i} & 2 & 2+2\mathbf{i} \\ -2\mathbf{i} & 0 & 0 & 2 \\ -2 & -2+2\mathbf{i} & 2+3\mathbf{i} & -2\mathbf{i} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

we have that $\mathrm{Sdet}(A) = \det(A^c) = 4 = \left[\mathrm{Ddet}(A)\right]^2$. □

1.2.3 The eigenvalue problem

As in the commutative case, eigenvalues play an important role in the solution of quaternionic differential and difference equations, that will be considered in subsequent chapters. However, since the multiplication of quaternions is noncommutative, a distinction between right and left eigenvalues must be made.

Definition 1.2.13. A quaternion λ is said to be a *right (left) eigenvalue* of $A \in \mathbb{H}^{n\times n}$ if $Av = v\lambda$ ($Av = \lambda v$), for some nonzero quaternionic vector $v \in \mathbb{H}^n$. The vector v is called *right (left) eigenvector* associated with λ. The set

$$\sigma_r(A) = \left\{\lambda \in \mathbb{H} : Av = v\lambda, \text{ for some } v \in \mathbb{H}^n \setminus \{0\}\right\}$$

is called the *right spectrum* of A. The *left spectrum* is similarly defined and denoted by $\sigma_l(A)$.

The *standard right eigenvalues* of the matrix $A \in \mathbb{H}^{n\times n}$ are the complex right eigenvalues of A with nonnegative imaginary part.

As shown in the next examples, there is no close relation between right and left eigenvalues.

Example 1.2.14. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{i} \end{bmatrix}.$$

It is clear that $\sigma_l(A) = \{1, \mathbf{i}\}$ and the eigenvectors associated with 1 and $\mathbf{i}$ are $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, respectively. On the other hand, all the elements of the equivalence class $[\mathbf{i}]$ as well as 1 are right eigenvalues. Therefore, $\sigma_l(A) \subset \sigma_r(A)$. $\square$

Example 1.2.15. The matrix

$$A = \begin{bmatrix} 0 & \mathbf{i} \\ \mathbf{j} & 0 \end{bmatrix}$$

has two left eigenvalues $\pm\frac{\sqrt{2}}{2}(\mathbf{i}+\mathbf{j})$ and infinitely many right eigenvalues which are all the quaternions satisfying $\lambda^4 + 1 = 0$, i.e., $\sigma_r(A) = \left[\frac{\sqrt{2}}{2}(-1+\mathbf{i})\right] \cup \left[\frac{\sqrt{2}}{2}(1+\mathbf{i})\right]$. Note that $\sigma_l(A) \cap \sigma_r(A) = \emptyset$. $\square$

Example 1.2.16. Let

$$A = \begin{bmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{bmatrix}.$$

The right eigenvalues of A are only 1 and -1. However, the left eigenvalues of A are all the elements of the form $-\mathbf{i}[\mathbf{i}]$, i.e., all the elements of the equivalence class $[\mathbf{i}]$, multiplied on the left by $-\mathbf{i}$, and thus $\sigma_r(A) \subset \sigma_l(A)$. $\square$

Right eigenvalues have been widely studied and are more useful for our purposes. In the sequel, we shall simply refer to them as *eigenvalues*.

The following propositions contain some results concerning eigenvalues of a quaternionic matrix as well as their relation with the eigenvalues of its complex adjoint matrix.

Proposition 1.2.17. [4, 55] *Let $A \in \mathbb{H}^{n\times n}$. Then*

$$\lambda \in \sigma_r(A) \Rightarrow [\lambda] \subseteq \sigma_r(A).$$

Proof. [55] Let $\lambda' \in [\lambda]$, i.e., there exists a nonzero $\alpha \in \mathbb{H}$ such that $\lambda' = \alpha\lambda\alpha^{-1}$. By hypothesis, $Av = v\lambda$, for some $v \in \mathbb{H}^n \setminus \{0\}$, which implies that $Av\alpha^{-1} = v\lambda\alpha^{-1}$. Thus, if we put $v' = v\alpha^{-1}$, we get that

$$Av' = v\alpha^{-1}\alpha\lambda\alpha^{-1} = v'\lambda'$$

which means that λ' is also a right eigenvalue of A. $\square$

Proposition 1.2.18. [55] *Let $A \in \mathbb{H}^{n\times n}$ and A^c be its complex adjoint matrix. Then*

$$\lambda \in \sigma(A^c) \Leftrightarrow F_A(\lambda) = 0 \quad \textit{and} \quad \lambda \in \sigma(A^c) \Rightarrow [\lambda] \subseteq \sigma_r(A).$$

Proposition 1.2.19. [30, 55] *Each matrix $A \in \mathbb{H}^{n\times n}$ has exactly $2n$ eigenvalues belonging to $\mathbb{C}$.*

Proof. [55] Let $A \in \mathbb{H}^{n\times n}$. Then, assuming that $\lambda \in \mathbb{C}$, the equation

$$Av = v\lambda,$$

is equivalent to, $(Av)^{\mathbb{C}} = (v\lambda)^{\mathbb{C}}$. But, by Proposition 1.2.3,

$$(Av)^{\mathbb{C}} = (v\lambda)^{\mathbb{C}} \Leftrightarrow A^c v^{\mathbb{C}} = v^c \lambda^{\mathbb{C}},$$

which yields $A^c v^{\mathbb{C}} = \lambda v^{\mathbb{C}}$. This implies that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if it is an eigenvalue of A^c. Taking into account that the size of A^c is $2n \times 2n$, we conclude that A has exactly $2n$ complex eigenvalues (counting with multiplicities). □

Proposition 1.2.20. [55] *Let $A \in \mathbb{H}^{n\times n}$ and A^c be its complex adjoint matrix. Then every real eigenvalue of A^c has even multiplicity and the complex eigenvalues of A^c appear in conjugate pairs.*

This proposition implies that the eigenvalues of a complex adjoint matrix can be divided into two sets such that any element of one set is the conjugate of an element of the other. Combining Propositions 1.2.19, 1.2.17 and 1.1.3 and the fact that $\eta \sim \overline{\eta}$, for all $\eta \in \mathbb{H}$, we obtain the following corollary.

Corollary 1.2.21. [30, 55] *Let $A \in \mathbb{H}^{n\times n}$. Then A has exactly n eigenvalues up to equivalence classes.*

In view of the previous results, one can compute the eigenvalues of a quaternionic matrix by means of the following simple procedure.

Let $A \in \mathbb{H}^{n\times n}$.

- Decompose $A = A_1 + A_2\mathbf{j}$, $A_1, A_2 \in \mathbb{C}^{n\times n}$;
- Construct the complex adjoint matrix $A^c = \begin{bmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{bmatrix}$;
- Calculate $\sigma(A^c) = \{\lambda_1, \ldots, \lambda_n, \overline{\lambda}_1, \ldots, \overline{\lambda}_n\}$;
- Divide $\sigma(A^c)$ into two sets σ_1, σ_2 such that $\sigma_1 \cup \sigma_2 = \sigma(A^c)$, the elements of σ_1 are the conjugates of the ones of σ_2 and the eigenvalues with positive imaginary part belong to σ_1;
- Then, $\sigma_r(A) = \bigcup_{p=1}^{n} [\lambda_p],\ \lambda_p \in \sigma_1$.

Note that the eigenvalues that belong to σ_1 are the *standard right eigenvalues* of A.

Example 1.2.22. Let

$$A = \begin{bmatrix} 0 & 0 \\ \mathbf{j}-\mathbf{k} & 1+\mathbf{i} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1+\mathbf{i} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1-\mathbf{i} & 0 \end{bmatrix}\mathbf{j}.$$

Then

$$A^c = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1+\mathbf{i} & 1-\mathbf{i} & 0 \\ 0 & 0 & 0 & 0 \\ -1-\mathbf{i} & 0 & 0 & 1-\mathbf{i} \end{bmatrix}$$

and

$$\sigma(A^c) = \{0, 1+\mathbf{i}, 0, 1-\mathbf{i}\}.$$

Therefore

$$\sigma_r(A) = \big\{0, [1+\mathbf{i}]\big\}.$$

$\square$

In the next example it is shown that, unlike the commutative case, the eigenvectors associated to different eigenvalues are not necessarily linearly independent.

Example 1.2.23. Let

$$A = \begin{bmatrix} \mathbf{i} & 1 \\ 0 & \mathbf{j} \end{bmatrix}.$$

Then the eigenvectors $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} \mathbf{i}+\mathbf{j} \\ 0 \end{smallmatrix}\right]$ associated with the eigenvalues $\mathbf{i}$ and $\mathbf{j}$, respectively, are not linearly independent. □

However, the following result holds.

Lemma 1.2.24. [55] *The eigenvectors associated with two eigenvalues are linearly independent if and only if the eigenvalues are not similar.*

Proposition 1.2.20 implies that if λ is an eigenvalue of A^c then its conjugate $\overline{\lambda}$ is also an eigenvalue of A^c. We next investigate what is the relation between generalized eigenvectors associated with conjugate eigenvalues.

Lemma 1.2.25. *Let $A \in \mathbb{H}^{n\times n}$. Then, for any v_1, v_2, w_1, $w_2 \in \mathbb{C}^n$*

$$A^c \begin{bmatrix} v_1 \\ -\overline{v}_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ -\overline{v}_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ -\overline{w}_2 \end{bmatrix} \Leftrightarrow A^c \begin{bmatrix} v_2 \\ \overline{v}_1 \end{bmatrix} = \overline{\lambda} \begin{bmatrix} v_2 \\ \overline{v}_1 \end{bmatrix} + \begin{bmatrix} w_2 \\ \overline{w}_1 \end{bmatrix}.$$

Proof. Let $A \in \mathbb{H}^{n\times n}$, $A^c = \begin{bmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{bmatrix}$ and v_1, v_2, w_1, $w_2 \in \mathbb{C}^n$. Then

$$\begin{aligned}
\begin{bmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{bmatrix} \begin{bmatrix} v_1 \\ -\overline{v}_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ -\overline{v}_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ -\overline{w}_2 \end{bmatrix} &\Leftrightarrow \begin{bmatrix} A_1 v_1 - A_2 \overline{v}_2 \\ -\overline{A}_2 v_1 - \overline{A}_1 \overline{v}_2 \end{bmatrix} = \begin{bmatrix} \lambda v_1 + w_1 \\ -\lambda \overline{v}_2 - \overline{w}_2 \end{bmatrix} \\
&\Leftrightarrow \begin{bmatrix} \overline{A}_1 \overline{v}_1 - \overline{A}_2 v_2 \\ A_2 \overline{v}_1 + A_1 v_2 \end{bmatrix} = \begin{bmatrix} \overline{\lambda} \overline{v}_1 + \overline{w}_1 \\ \overline{\lambda} v_2 + w_2 \end{bmatrix} \\
&\Leftrightarrow \begin{bmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{bmatrix} \begin{bmatrix} v_2 \\ \overline{v}_1 \end{bmatrix} = \overline{\lambda} \begin{bmatrix} v_2 \\ \overline{v}_1 \end{bmatrix} + \begin{bmatrix} w_2 \\ \overline{w}_1 \end{bmatrix}
\end{aligned}$$

□

If $w_1 = w_2 = 0$ this means that if $\left[\begin{smallmatrix} v_1 \\ -\overline{v}_2 \end{smallmatrix}\right]$ is an eigenvector of A^c associated with λ then $\left[\begin{smallmatrix} v_2 \\ \overline{v}_1 \end{smallmatrix}\right]$ is an eigenvector of A^c associated with $\overline{\lambda}$. The successive application of the Lemma allows to conclude that if $\left[\begin{smallmatrix} v_1 \\ -\overline{v}_2 \end{smallmatrix}\right]$ is a generalized eigenvector of A^c associated with λ then $\left[\begin{smallmatrix} v_2 \\ \overline{v}_1 \end{smallmatrix}\right]$ is a generalized eigenvector of A^c associated with $\overline{\lambda}$.

1.3 An application of quaternions

The rotation of a vector in $\mathbb{R}^3$ is done by multiplying it by a 3×3 unitary matrix. Such matrix can be constructed knowing the axis and the angle of rotation or using the so called Euler angles which define a rotation as a composition of three independent rotations about the coordinate axes. Alternatively, rotations can also be achieved using unitary quaternions (which lie on a 4-dimensional sphere). It turns out that this method has some advantages, for example in computational efficiency. Indeed, a quaternion is represented by 4 real numbers whereas a 3×3 matrix requires 9 numbers. Moreover, the composition of two rotations in a quaternionic framework requires 16 multiplications and 12 additions, while the same operation using 3×3 matrices requires 27 multiplications and 18 additions [33]. Another advantage is given by the possibility of doing interpolation. This technique is mostly applied in computer graphics where, for instance in computer games, the view of a scene while rotating the camera should be as smooth as possible. This can be easily obtained by using quaternions since the interpolation of two quaternions can be done uniformly along a geodesic on the surface of the 4-D sphere [46]. One of the first computer games to use quaternions in their animation to rotate objects has been the famous “Tomb Raider”.

In this section, following [26], we show how quaternions can be used to describe rotations in a three dimensional space.

Let $u = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ and $v = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$ be vectors in $\mathbb{R}^3$. Suppose that the rotation of v by an angle θ about the direction of u yields the vector $\widetilde{v} = \begin{bmatrix} \widetilde{v_1} & \widetilde{v_2} & \widetilde{v_3} \end{bmatrix}^T$ as illustrated in Figure 1.

Identify v and u with the pure quaternions $v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, respectively, and define the unit quaternion

$$q = \cos\frac{\theta}{2} + \frac{u}{|u|}\sin\frac{\theta}{2}.$$

Then, it can be shown that $qv\overline{q} = \widetilde{v}_1\mathbf{i} + \widetilde{v}_2\mathbf{j} + \widetilde{v}_3\mathbf{k}$.

Note that, for any $q, v \in \mathbb{H}$, the quaternion $qv\overline{q}$ has zero real part whenever so has v and so the following theorem holds.

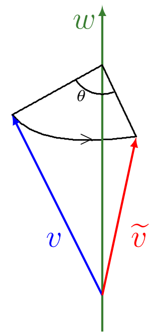

Figure 1

Theorem 1.3.1. [26] *For any unit quaternion*

$$q = \cos\frac{\theta}{2} + w\sin\frac{\theta}{2}$$

and for any vector $v \in \mathbb{R}^3$ the action of the operator $L_q(v) = qv\overline{q}$ consists in a rotation of v by an angle θ about the direction specified by the vector w.

If q is not unitary, then the operator L_q defined in Theorem 1.3.1 acts as a rotation composed with a norm variation. So, to obtain only a rotation, when q is not unitary the operator $\widetilde{L}_q : v \mapsto qvq^{-1}$ is used, which is well defined for any $q \neq 0$. Indeed, it is easy to check that $\widetilde{L}_q = L_{\frac{q}{|q|}}$.

Chapter 2

Quaternionic polynomials

Similar to what happens in the real and complex cases, polynomials an polynomial matrices are an important tool in the analysis of the dynamical properties of quaternionic behavioral systems.

In this chapter, after introducing the quaternionic polynomials, some basic notions, such as zeros and divisors, are given and characterized. It turns out that the relation between zeros and divisors is not as simple as in the commutative case, which leads to some surprising results. To conclude the chapter we consider a new similarity relation that will play an important role in the sequel.

2.1 Definitions and basic properties

Unlike the real or complex case, there are several possible ways to define quaternionic polynomials since the coefficients can be taken to be on the right, on the left or on both sides of the indeterminate (see, e.g., [44]). Here we take the coefficients to be on the left of the indeterminate. This choice will become clear in Chapter 4, where such polynomials are associated with linear difference and differential equations with quaternionic coefficients.

Definition 2.1.1. A *quaternionic Laurent-polynomial* (or *L-polynomial*) $p(s, s^{-1})$ is defined by

$$p(s, s^{-1}) = \sum_{l=M}^{N} p_l s^l$$

where $p_l \in \mathbb{H},\ p_N \neq 0 \neq p_M,\ M, N \in \mathbb{Z}, M \leq N$.

The *degree* of $p(s, s^{-1})$, $\deg p(s, s^{-1})$, is $N-M$ and its *leading coefficient*, $\operatorname{lc} p(s, s^{-1})$, is p_N. As usual, if $\operatorname{lc} p(s, s^{-1}) = 1$, $p(s, s^{-1})$ is said to be *monic*. If $M \geq 0$, $p(s, s^{-1})$ is simply said to be a *quaternionic polynomial* and we denote it by $p(s)$. The degree of the polynomial $p(s)$ is N.

The sets of quaternionic Laurent-polynomials and quaternionic polynomials are denoted by $\mathbb{H}[s, s^{-1}]$ and $\mathbb{H}[s]$, respectively.

Although later in this thesis L-polynomials are more relevant than polynomials, for the sake of simplicity, in the sequel our definitions and results are stated only for quaternionic polynomials, since they trivially generalized to L-polynomials.

The set of quaternionic polynomials can be endowed with two binary operations, sum and product, which are defined as in the commutative case with the additional rule

$$(\alpha s^n)(\beta s^m) = \alpha\beta s^{n+m},\ \ \alpha, \beta \in \mathbb{H},$$

i.e., the monomial product is performed as if the indeterminate *commuted* with the coefficients. Indeed, given two quaternionic polynomials $p(s) = \sum_{n=0}^{N} p_n s^n$ and $q(s) = \sum_{m=0}^{M} q_m s^m$, their sum and product are defined, respectively, as

$$p(s) + q(s) = \sum_{l=0}^{\max\{N,M\}} (p_l + q_l) s^l,$$

$$p(s)q(s) = \sum_{l=0}^{N+M} \sum_{n+m=l} p_n q_m s^l.$$

It is clear that $\mathbb{H}[s]$ endowed with the defined operations is a noncommutative ring [23].

Given a polynomial $r(s) = \sum_{n=0}^{N} r_n s^n$ and an element $\lambda \in \mathbb{H}$, the *evaluation* of $r(s)$ in λ, $r(\lambda)$, is defined to be the element

$$r(\lambda) = \sum_{n=0}^{N} r_n \lambda^n \in \mathbb{H}.$$

Note that, unlike the commutative case, evaluation of polynomials is not a ring homomorphism, i.e., if $r(s) = p(s)q(s) \in \mathbb{H}[s]$, then in general $r(\lambda) \neq p(\lambda)q(\lambda)$, $\lambda \in \mathbb{H}$, as shown in the following example. It is clear however that if $\lambda \in \mathbb{R}$, $r(\lambda) = p(\lambda)q(\lambda)$.

Example 2.1.2. Consider the quaternionic polynomials $p(s) = s - \mathbf{i}$ and $q(s) = s - \mathbf{j}$ and let

$$r(s) = p(s)q(s) = (s - \mathbf{i})(s - \mathbf{j}) = s^2 - (\mathbf{i} + \mathbf{j})s + \mathbf{k}.$$

Then

$$r(\mathbf{i}) = \mathbf{i}^2 - (\mathbf{i} + \mathbf{j})\mathbf{i} - \mathbf{k} = 2\mathbf{k} \quad \text{but} \quad p(\mathbf{i})q(\mathbf{i}) = (\mathbf{i} - \mathbf{i})(\mathbf{i} - \mathbf{j}) = 0.$$

$\square$

To simplify the notation, we will indicate the product of polynomials $p(s)$ and $q(s)$ as $pq(s)$. We will also omit the indeterminate s and write $p \in \mathbb{H}[s]$ if no ambiguity arises.

A polynomial $p(s)$ is said to be an *invertible element* or a *unit* of $\mathbb{H}[s]$ if there exists $q(s) \in \mathbb{H}[s]$ such that $q(s)p(s) = p(s)q(s) = 1$. Clearly, the units of $\mathbb{H}[s]$ are only the constant monomials $u(s) = \alpha,\ \alpha \in \mathbb{H}\backslash\{0\}$. Note that, however, the units of $\mathbb{H}[s, s^{-1}]$ are the monomials $\alpha s^m,\ m \in \mathbb{Z},\ \alpha \in \mathbb{H}\backslash\{0\}$.

Conjugacy is extended to quaternionic polynomials by linearity and according to the rule $\overline{\alpha s^n} = \overline{\alpha} s^n,\ \forall \alpha \in \mathbb{H}$, i.e., the *conjugate* of $p(s) = \sum_{l=0}^{N} p_l s^l \in \mathbb{H}[s]$ is

$$\overline{p}(s) = \sum_{l=0}^{N} \overline{p}_l s^l.$$

Properties related to conjugation of quaternions extend to polynomials as shown in the following proposition.

Proposition 2.1.3. [27] *Let $p, q \in \mathbb{H}[s]$. Then*

1. $\overline{pq} = \overline{q}\,\overline{p}$.
2. $p\overline{p} = \overline{p}p \in \mathbb{R}[s]$.
3. *If $pq \in \mathbb{R}[s]$, then $pq = qp$.*

Proof. 1. This follows trivially from the definition of product of quaternionic polynomials, taking into account that $\overline{\eta\nu} = \overline{\nu}\,\overline{\eta}$, $\forall \eta, \nu \in \mathbb{H}$.

2. By 1., we have that

$$\overline{\overline{p}p} = \overline{p}\,\overline{\overline{p}} = \overline{p}p$$

which implies that $\overline{p}p \in \mathbb{R}[s]$. Therefore $\overline{p}p$ commutes with p, i.e., $p\overline{p}p = \overline{p}pp$. Hence $p\overline{p} = \overline{p}p$.

3. Real polynomials commute with any polynomial. Thus, by 2.,

$$\overline{q}pq = pq\overline{q} = q\overline{q}p = \overline{q}qp,$$

and so, $pq = qp$. □

2.2 Divisors and zeros

In this section the notions of divisor and zero of a quaternionic polynomial are introduced and the relation between zeros and divisors is analyzed. Although this subject is well documented in the literature, see for instance [3, 16, 27, 36], we have opted to present it here in detail due to its fundamental importance throughout this thesis. After this overview of known results, a new definition of total divisor is presented, which proves to be equivalent to other existing characterizations.

2.2.1 Divisors

Since the product of quaternionic polynomials is noncommutative, it is necessary to make a distinction between *left* and *right* multiplication and related notions.

The polynomial ring $\mathbb{H}[s]$ is *right-* and *left-Euclidean*, i.e., it allows *right* and *left division with remainder*. Actually, for every two quaternionic polynomials $p(s)$ and $d(s)$, with $d(s)$ nonzero, there exist quaternionic polynomials $q_1(s)$, $r_1(s)$, $q_2(s)$, $r_2(s)$ such that

$$p(s) = q_1(s)d(s) + r_1(s) \quad \text{with} \ \deg r_1(s) < \deg d(s) \quad \text{or} \quad r_1(s) = 0$$

for the right division, and

$$p(s) = d(s)q_2(s) + r_2(s) \quad \text{with} \ \deg r_2(s) < \deg d(s) \quad \text{or} \quad r_2(s) = 0$$

for the left division.

Example 2.2.1. Take $p(s) = s^4 + (\mathbf{i} + \mathbf{j})s^3 + \mathbf{k}s^2 + 1$ and $d(s) = s^2 + \mathbf{j}s$. Then

$$p(s) = \left(s^2 + \mathbf{i}s\right) d(s) + 1 \quad \text{and} \quad p(s) = d(s) \left(s^2 + \mathbf{i}s + 2\mathbf{k}\right) - 2\mathbf{i}s + 1.$$

Note that the degrees of the left and right remainders are different. Indeed, $\deg(1) = 0$ while $\deg(-2\mathbf{i}s + 1) = 1$. □

A quaternionic polynomial $d(s)$ is a *left divisor* of a polynomial $p(s) \in \mathbb{H}[s]$, which we will denote by $d(s) \,|_l\, p(s)$, and $p(s)$ is a *right multiple* of $d(s)$, if there exists a polynomial $q(s)$ such that

$$p(s) = d(s)q(s).$$

If $d(s)$ is a left divisor of both $p(s)$ and $q(s)$, and $d(s)$ is a right multiple of every common left divisor of $p(s)$ and $q(s)$, then $d(s)$ is a *greatest common left divisor* (gcld) of $p(s)$ and $q(s)$. It is easily shown that the gcld is unique up to right multiplication by a unit. Two polynomials $p(s)$ and $q(s)$ are *left coprime* if every gcld of $p(s)$ and

$q(s)$ is a unit.

The definitions of *right divisor (left multiple)*, *greatest common right divisor* (gcrd), and *right coprimeness* are entirely analogous. We will use the notation $d(s) \,|_r\, p(s)$ to indicate that $d(s)$ is a right divisor of $p(s)$.

A polynomial $d(s)$ is a *divisor* of a polynomial $p(s)$, which we will denote by $d(s) \,|\, p(s)$, and $p(s)$ is a *multiple* of $d(s)$, if $d(s) \,|_l\, p(s)$ and $d(s) \,|_r\, p(s)$.

In general, the gcld's of two quaternionic polynomials are different from their gcrd's.

Example 2.2.2. Let $p(s) = \mathbf{j}s - \mathbf{k}$ and $q(s) = -\mathbf{i}s + 1$. Then every gcrd $\big(p(s), q(s)\big)$ is of the form $\eta(-\mathbf{i}s + 1)$, $\eta \in \mathbb{H}$, and every gcld $\big(p(s), q(s)\big)$ is a nonzero constant $\nu \in \mathbb{H} \setminus \{0\}$.

Thus p and q are left coprime, but not right coprime. □

2.2.2 Zeros

The *zeros* of a polynomial $p \in \mathbb{H}[s]$ are the values $\lambda \in \mathbb{H}$ such that $p(\lambda) = 0$. A pair $(p, q) \in \mathbb{H}[s] \times \mathbb{H}[s]$ is *zero coprime* if p and q do not have common zeros.

Factors of a polynomial are usually related to its zeros, but the fact that, as mentioned before, evaluation is not a ring homomorphism implies that the relation between the factors and the zeros of a quaternionic polynomial is not as simple as for real or complex polynomials.

The next proposition establishes a connection between zeros and right divisors.

Proposition 2.2.3. [27] *A quaternion α is a zero of a nonzero $p \in \mathbb{H}[s]$ if and only if the polynomial $s - \alpha$ is a right divisor of p.*

Proof. "If" part. Assume that $s - \alpha$ is a right divisor of p. Then there exist $p_l \in \mathbb{H}$,

$l = 0, \ldots, n$, such that

$$p(s) = \left(\sum_{l=0}^{n} p_l s^l\right)(s-\alpha) = \sum_{l=0}^{n} p_l s^{l+1} - \sum_{l=0}^{n} p_l \alpha s^l,$$

and therefore

$$p(\alpha) = \sum_{l=0}^{n} p_l \alpha^{l+1} - \sum_{l=0}^{n} p_l \alpha \alpha^l = 0.$$

"Only if" part. Assume that $p(\alpha) = 0$. As mentioned before, by the right-Euclidean algorithm there exist $d \in \mathbb{H}[s]$ and $\beta \in \mathbb{H}$ such that $p(s) = d(s)(s-\alpha) + \beta$. From the first part we have that α is a zero of the polynomial $d(s)(s-\alpha)$ and thus $0 = p(\alpha) = \beta$, i.e., $p(s) = d(s)(s-\alpha)$. $\square$

Note that this implies that zero coprimeness is equivalent to right coprimeness. However, if $d(s)$ is a left divisor of $p(s)$, the zeros of $d(s)$ are not necessarily zeros of $p(s)$ as the following example shows.

Example 2.2.4. Let $d(s) = s - \mathbf{i}$ and $p(s) = d(s)\mathbf{j} = \mathbf{j}s - \mathbf{k}$. Then

$$d(\mathbf{i}) = 0 \quad \text{but} \quad p(\mathbf{i}) = \mathbf{ji} - \mathbf{k} = -2\mathbf{k} \neq 0.$$

$\square$

Nevertheless, there is still some connection between the zeros of a polynomial and the zeros of its left divisors. Indeed, let $r = pq \in \mathbb{H}[s]$, if $\alpha \in \mathbb{H}$ is a zero of the polynomial r but not of its right divisor q, then its left divisor p must have a zero that is equivalent to α. This is formalized in the following result.

Proposition 2.2.5. [27] *Let $r = pq \in \mathbb{H}[s]$ and $\alpha \in \mathbb{H}$ be such that $\beta = q(\alpha) \neq 0$. Then*

$$r(\alpha) = p\left(\beta\alpha\beta^{-1}\right)q(\alpha).$$

In particular, α is a zero of r if and only if $\beta\alpha\beta^{-1}$ is a zero of p.

Proof. Let $p(s) = \sum p_l s^l$. Then $r(s) = p(s)q(s) = \left(\sum p_l s^l\right) q(s) = \sum p_l q(s) s^l$, and so

$$\begin{aligned} r(\alpha) &= \sum p_l q(\alpha)\alpha^l = \sum p_l \beta \alpha^l = \sum p_l \beta \alpha^l \beta^{-1} \beta \\ &= \sum p_l \left(\beta\alpha\beta^{-1}\right)^l \beta = p\left(\beta\alpha\beta^{-1}\right) q(\alpha). \end{aligned}$$

If $r(\alpha) = 0$ and $\beta = q(\alpha) \neq 0$, then

$$r(\alpha) = p\left(\beta\alpha\beta^{-1}\right) q(\alpha) = 0.$$

Since $\mathbb{H}[s]$ has no zero-divisors it follows that $\beta\alpha\beta^{-1}$ is a zero of p. On the other hand, it is obvious that if $p\left(\beta\alpha\beta^{-1}\right) = 0$ then $r(\alpha) = 0$. $\square$

As in the commutative case, every quaternionic polynomial can be decomposed in linear factors, i.e., for each $p \in \mathbb{H}[s]$ of degree n there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{H}$ such that $p = (s - \alpha_1) \cdots (s - \alpha_n)$ [27, Theorem 16.9]. However, this factorization is not unique. Indeed, the quaternionic polynomial $p = s^2 + 1$ can be factorized either as $p = (s - \mathbf{i})(s + \mathbf{i})$ or as $p = (s - \mathbf{j})(s + \mathbf{j})$.

Nevertheless, there exists a relation between all the possible factorizations. For instance, let $(s - \alpha_1) \cdots (s - \alpha_n)$ and $(s - \alpha_1') \cdots (s - \alpha_n')$ be two factorizations of a quaternionic polynomial. Then, each α_l, $l = 1, \ldots, n$, is similar to some α_m', $m = 1, \ldots, n$, and vice-versa. Moreover, the following proposition holds.

Proposition 2.2.6. [16, 27] *If $p = (s - \alpha_1) \cdots (s - \alpha_n) \in \mathbb{H}[s]$, where $\alpha_1, \ldots, \alpha_n \in \mathbb{H}$, then every zero of p is similar to some α_l, $l = 1, \ldots, n$. Reciprocally, every α_l is similar to some zero of p.*

Remark 2.2.7. Note that, as a consequence of Proposition 2.2.3, α_n is itself a zero of $p = (s - \alpha_1) \cdots (s - \alpha_n) \in \mathbb{H}[s]$, since it is a zero of its right divisor $s - \alpha_n$. $\square$

Corollary 2.2.8. *Let $p \in \mathbb{H}[s]$. If $p(\alpha) = 0$ for some $\alpha \in \mathbb{H}$, then there exists $\alpha' \in [\alpha]$ such that $\overline{p}(\alpha') = 0$.*

Proof. Since $p(\alpha) = 0$, by Proposition 2.2.6, α is similar to some α'' such that $s - \alpha''$ is a factor of p. By Proposition 2.1.3-1, $s - \overline{\alpha''}$ is a factor of $\overline{p}$, thus, again by Proposition 2.2.6, $\overline{\alpha''}$ is similar to some zero α' of $\overline{p}$. Since $\overline{\alpha''} \sim \alpha''$, we conclude that $\alpha' \in [\alpha''] = [\alpha]$, yielding the desired result. $\square$

It is well know that, over the real or complex numbers, a polynomial of degree n has at most n distinct zeros. For quaternionic polynomials this is no longer true; for instance, it is clear that the imaginary units **i**, **j** and **k** are all zeros of the polynomial $p(s) = s^2 + 1 \in \mathbb{H}[s]$. Still, the following proposition holds.

Proposition 2.2.9. [16, 27] *Let $p \in \mathbb{H}[s]$ be a polynomial of degree n. Then*

1. *The zeros of p belong at most to n equivalence classes.*

2. *If p has two distinct zeros in an equivalence class, then it has infinitely many zeros in that equivalence class.*

3. *The number of zeros of p is not larger than n or infinite.*

The next lemma collects some basic results about zeros of quaternionic polynomials which can be found in [3]. Before stating it, we first define the *minimal polynomial of the equivalence class* $[\lambda]$ as the real irreducible monic polynomial

$$\psi_{[\lambda]} = (s - \overline{\lambda})(s - \lambda) = s^2 - 2(\operatorname{Re}\lambda)s + |\lambda|^2 \tag{2.1}$$

if $\lambda \notin \mathbb{R}$ and as $\psi_{[\lambda]} = s - \lambda$ if $\lambda \in \mathbb{R}$.

Lemma 2.2.10. [3] *Let $p \in \mathbb{H}[s]$. Then*

1. *$\psi_{[\lambda]} = \psi_{[\lambda']}$ if and only if $[\lambda] = [\lambda']$ (i.e., $\psi_{[\lambda]}$ is well-defined).*

2. *If $\psi_{[\lambda]} \,|\, p$ then $p(\nu) = 0$ for every $\nu \in [\lambda]$. Otherwise, at most one $\nu \in [\lambda]$ is a zero of p.*

3. *If $p(\nu) = p(\lambda) = 0$ with $\nu \neq \lambda$, $\nu \in [\lambda]$, then $\psi_{[\lambda]} \,|\, p$.*

4. *Suppose that $p(\lambda) = 0$. Then $\psi_{[\lambda]} \,|\, \overline{p}p$. If in particular $p \in \mathbb{R}[s]$, $\psi_{[\lambda]} \,|\, p$.*

5. *If $\psi_{[\lambda]} \,|\, \overline{p}p$ then there exists $\lambda' \in [\lambda]$ such that $p(\lambda') = 0$.*

Proof. 1. Simply note that, by its definition, $\psi_{[\lambda]}$ depends only on $\operatorname{Re}\lambda$ and on $|\lambda|^2$ that uniquely characterize and are characterized by the equivalence class $[\lambda]$ (see Proposition 1.1.2).

2. If $\lambda \in \mathbb{R}$ the result is immediate. Suppose first that $\psi_{[\lambda]} \,|\, p$ and let $\nu \in [\lambda]$. By 1., also $\psi_{[\nu]} \,|\, p$, thus the polynomial $s - \nu$ is a right divisor of p and hence, by Proposition 2.2.3, $p(\nu) = 0$. Now suppose that $\psi_{[\lambda]}$ does not divide p, i.e., there exist $d \in \mathbb{H}[s]$ and $\alpha, \beta \in \mathbb{H}$ non simultaneously zero, such that $p(s) = d(s)\psi_{[\lambda]}(s) + \alpha s + \beta$. By 1., for every $\nu \in [\lambda]$, $p(\nu) = \alpha\nu + \beta$. If $\alpha = 0$, then $p(\nu) = \beta \neq 0$ and the result follows. Finally, if $\alpha \neq 0$, then $p(\nu) = 0$ implies that $\nu = -\alpha^{-1}\beta$, and therefore there is a zero of p similar to λ if $-\alpha^{-1}\beta \in [\lambda]$ and none otherwise.

3. This is a consequence of 2.

4. Because $p(\lambda) = 0$, $s - \lambda$ is a right divisor of p, i.e., $p = d(s - \lambda)$ for some $d \in \mathbb{H}[s]$. Taking into account that $p(s)$ is a right factor of $\overline{p}p(s)$, also $\overline{p}p(\lambda) = 0$. Now, if $\lambda \in \mathbb{R}$ then $\psi_{[\lambda]} = s - \lambda$ and $\overline{p} = \overline{d}(s - \lambda)$. Hence, $\overline{p}p = \overline{d}d(s - \lambda)^2$ and the result follows. If $\lambda \notin \mathbb{R}$, as the polynomial $\overline{p}p$ is real by Proposition 2.1.3-2, $\overline{p}p(\overline{\lambda}) = \overline{\overline{p}p(\lambda)} = 0$. Since $\overline{\lambda} \in [\lambda]$, by 3., the desired result follows. The conclusion for the case when $p \in \mathbb{R}[s]$ is obtained analogously.

5. If $p(\lambda) = 0$ the statement is obviously true. If $p(\lambda) \neq 0$, by Proposition 2.2.5 there exists $\lambda^\star \in [\lambda]$ such that $0 = \overline{p}p(\lambda) = \overline{p}(\lambda^\star)p(\lambda)$. This implies that $\overline{p}(\lambda^\star) = 0$. But by Corollary 2.2.8 there exists $\lambda' \in [\lambda^\star] = [\lambda]$ such that $p(\lambda') = 0$. □

As mentioned earlier, a quaternionic polynomial may have different factorizations. However, if $p \in \mathbb{H}[s]$ only has a zero, it turns out that its factorization is unique. Indeed, it is clear that $\lambda \in \mathbb{R}$ is the unique zero of $p \in \mathbb{H}[s]$ with $\deg p = n$ if and only if p only admits the factorization $p(s) = (s - \lambda)^n$. If $\lambda \notin \mathbb{R}$ the following result holds.

Lemma 2.2.11. *Let $p \in \mathbb{H}[s]$ have degree n. Then $\lambda_1 \in \mathbb{H} \setminus \mathbb{R}$ is the unique zero of p if and only if p admits a unique factorization which has the form*

$$p(s) = (s - \lambda_n) \cdots (s - \lambda_2)(s - \lambda_1), \ \lambda_l \in [\lambda_1], \ \lambda_l \neq \overline{\lambda_{l-1}}, \ l = 2, \ldots, n. \qquad (2.2)$$

Proof. "If" part. It is obvious by Proposition 2.2.3.

"Only if" part. Let $\lambda_1 \in \mathbb{H} \setminus \mathbb{R}$ be the unique zero of p. By Proposition 2.2.3, $(s - \lambda_1)$ is the unique right factor of p and hence, every factorization of p has the form

$$p(s) = (s - \lambda_n) \cdots (s - \lambda_2)(s - \lambda_1), \qquad \text{for some } \lambda_2, \ldots, \lambda_n \in \mathbb{H}.$$

By Proposition 2.2.6 we have that $\lambda_l \in [\lambda_1],\ l = 2, \ldots, n$. Moreover, if by contradiction $\lambda_l = \overline{\lambda_{l-1}}$ for some l, then $\psi_{[\lambda_1]} \,|\, p$ and, by Lemma 2.2.10-2, this would imply that $p(\nu) = 0$ for every $\nu \in [\lambda_1]$, which is impossible since λ_1 is the unique zero of p. Therefore all the factorizations of p have the form (2.2). We show next that the factorization is unique. Suppose that

$$p(s) = (s - \lambda_n) \cdots (s - \lambda_2)(s - \lambda_1) = (s - \lambda_n') \cdots (s - \lambda_2')(s - \lambda_1), \tag{2.3}$$

are two factorizations of p, where $\lambda_l \in [\lambda_1],\ \lambda_l \neq \overline{\lambda_{l-1}}$ and $\lambda_l' \in [\lambda_1],\ \lambda_l' \neq \overline{\lambda_{l-1}'}$. After dividing $p(s)$ on the right by $(s - \lambda_1)$, from equation (2.3) we obtain

$$p_1(s) = (s - \lambda_n) \cdots (s - \lambda_2) = (s - \lambda_n') \cdots (s - \lambda_2').$$

Note that $\lambda_2' \in [\lambda_2] = [\lambda_1]$ and so, if by contradiction, $\lambda_2 \neq \lambda_2'$ then, by Lemma 2.2.10-3, $\psi_{[\lambda_2]} \,|\, p_1$. Thus $\psi_{[\lambda_2]} \,|\, p_1 \,|_l\, p$, which is impossible by the same arguments used before. Hence $\lambda_2 = \lambda_2'$. By repeating the same procedure it can be proved that $\lambda_l = \lambda_l'$, $l = 2, \ldots, n$, and therefore p admits a unique factorization. $\square$

Another issue that is not as immediate as for real polynomials is the construction of a quaternionic polynomial with prescribed zeros. In [6, 10], algorithms to find such polynomials are given. For instance, if two quaternions α and $\beta \notin [\alpha]$ are prescribed, the polynomial

$$p(s) = \big(s - (\beta - \alpha)\beta(\beta - \alpha)^{-1}\big)(s - \alpha)$$

has precisely α and β as zeros. If α and $\beta \in [\alpha] \setminus \{\alpha\}$ are prescribed, by Lemma 2.2.10-3 the only monic second degree polynomial that has α and β as zeros is $\psi_{[\alpha]}$.

On the other hand, the zeros of the quaternionic polynomial $q(s) = (s - \beta)(s - \alpha)$, where $\alpha, \beta \in \mathbb{H}$ and $\alpha \neq \overline{\beta}$, are

$$\alpha \quad \text{and} \quad \widetilde{\beta} = (\overline{\beta} - \alpha)^{-1}\beta(\overline{\beta} - \alpha).$$

Note that, this is in accordance with Proposition 2.2.6, since the quaternionic polynomial q has indeed a zero, $\widetilde{\beta}$, within the equivalence class of β. If $\alpha = \overline{\beta}$, then the zeros of q are all the quaternions similar to β since, in this case, $q(s) = (s-\beta)(s-\overline{\beta}) = \psi_{[\beta]}$.

Example 2.2.12. 1. The quaternionic polynomial q such that $q(\mathbf{i}) = q(2\mathbf{k}) = 0$ is

$$\begin{aligned} q(s) &= \big(s - (2\mathbf{k}-\mathbf{i})2\mathbf{k}(2\mathbf{k}-\mathbf{i})^{-1}\big)(s-\mathbf{i}) \\ &= \left(s - \tfrac{6\mathbf{k}-8\mathbf{i}}{5}\right)(s-\mathbf{i}) \\ &= s^2 - \tfrac{6\mathbf{k}-3\mathbf{i}}{5}s + \tfrac{6\mathbf{j}+8}{5}. \end{aligned}$$

2. The quaternionic polynomial $q(s) = (s-\mathbf{j})(s-\mathbf{i})$ has a unique zero, $\mathbf{i}$, since

$$(-\mathbf{j}-\mathbf{i})^{-1}\mathbf{j}(-\mathbf{j}-\mathbf{i}) = \mathbf{i}.$$

□

As is well-known, in the commutative case the factors of a polynomial are related to its zeros and the corresponding multiplicities. Since in the quaternionic case the factorization of a polynomial is not unique, there is no natural way to define multiplicity of a zero of a quaternionic polynomial. The definition that we give here is related to right factors as follows. The *multiplicity* of λ as a zero of $p \in \mathbb{H}[s]$, $\mu_\lambda(p)$, is the maximum degree of the right factors of p having λ as their unique zero. Note that if $p(\lambda) \neq 0$, $\mu_\lambda(p) = 0$.

Example 2.2.13. The polynomial $\psi_{[\mathbf{i}]}(s) = s^2+1$ has infinitely many zeros, $\lambda \in [\mathbf{i}]$, all with multiplicity one. On the other hand, both

$$p(s) = (s-\mathbf{k})(s-\mathbf{i}) \qquad \text{and} \qquad q(s) = (s-\mathbf{j})(s-\mathbf{i})$$

have a unique zero, $\lambda = \mathbf{i}$, (see Example 2.2.12-2), and therefore $\mu_{\mathbf{i}}(p) = \mu_{\mathbf{i}}(q) = 2$. □

Remark 2.2.14. Let $p \in \mathbb{H}[s]$, $\lambda \notin \mathbb{R}$ and $\mu_\lambda(p) = n$. Then, by definition, p has a right divisor $d \in \mathbb{H}[s]$ of degree n whose unique zero is λ. By Lemma 2.2.11, d has the form

$$d(s) = (s-\lambda_n)\cdots(s-\lambda_2)(s-\lambda), \quad \lambda_l \in [\lambda],\ l = 2,\ldots,n.$$

Therefore, $d\overline{d} = \psi^n_{[\lambda]}$. □

In the next lemma we present two results on multiplicity of λ as a zero of $\psi_{[\lambda]}$ that will be used in the following section.

Lemma 2.2.15. *Let $p \in \mathbb{H}[s]$ and $\lambda \in \mathbb{H}$. Then*

1. $\mu_\lambda\left(\psi^n_{[\lambda]}\right) = n$.
2. $\mu_\lambda\left(\psi^n_{[\lambda]}p\right) = \mu_\lambda(p) + n$.

Proof. 1. If $\lambda \in \mathbb{R}$ the result is trivial. If $\lambda \in \mathbb{H}\setminus\mathbb{R}$, since $\psi_{[\lambda]} \in \mathbb{R}[s]$, as a consequence of Proposition 2.1.3-3 we have that

$$\psi^n_{[\lambda]} = \left[(s-\overline{\lambda})(s-\lambda)\right]^n = (s-\overline{\lambda})^n(s-\lambda)^n.$$

By Lemma 2.2.11, it is clear that $(s-\lambda)^n$ is the maximum degree right factor of $\psi^n_{[\lambda]}$ having λ as it unique zero and thus $\mu_\lambda\left(\psi^n_{[\lambda]}\right) = n$.

2. If $\lambda \in \mathbb{R}$ the result is obvious. Suppose that $\lambda \in \mathbb{H}\setminus\mathbb{R}$ and let $\mu_\lambda(p) = m$. Then p has a right divisor $d \in \mathbb{H}[s]$ of degree m whose unique zero is λ. By Lemma 2.2.11, d can be factorized as

$$d(s) = (s-\lambda_m)\cdots(s-\lambda_2)(s-\lambda), \quad \lambda_l \in [\lambda],\ l = 2,\ldots,m.$$

Therefore, $p = rd$, for some $r \in \mathbb{H}[s]$, and $\psi^n_{[\lambda]}p = r\psi^n_{[\lambda]}d$. Moreover, since $\lambda_m \in [\lambda]$, $\psi^n_{[\lambda]} = (s-\overline{\lambda_m})^n(s-\lambda_m)^n$ and hence $\psi^n_{[\lambda]}p = r'd'$ with

$$r'(s) = r(s)(s-\overline{\lambda_m})^n \qquad \text{and} \qquad d'(s) = (s-\lambda_m)^n d(s).$$

By Lemma 2.2.11, λ is the unique zero of d' and therefore $\mu_\lambda(d') = m+n$. Clearly, d' is the maximum degree right factor of $\psi^n_{[\lambda]}p$ having λ as it unique zero and thus $\mu_\lambda(\psi^n_{[\lambda]}p) = m+n = \mu_\lambda(p)+n$. $\square$

2.2.3 Total divisor

In the real or complex case, the divisibility property is enough to define the Smith form of a polynomial matrix. However, it turns out that in order to define the Smith and

Smith-McMillan forms in the quaternionic case, that will be given in the next chapter, a stronger concept of divisibility has to be introduced. First it is necessary to extend the similarity relation to quaternionic polynomials.

Two quaternionic polynomials $p(s)$ and $q(s)$ are said to be *similar*, $p(s) \sim q(s)$, if there exists a nonzero $\alpha \in \mathbb{H}$ such that $p(s) = \alpha q(s)\alpha^{-1}$. Clearly, this is an equivalence relation. We denote by $[q(s)]$ the equivalence class containing $q(s)$.

Remark 2.2.16. Note that if the quaternionic polynomials $p(s) = \sum_{n=0}^{N} p_n s^n$ and $q(s) = \sum_{n=0}^{N} q_n s^n$ are similar then $p_l \sim q_l$, for all $l = 0, \ldots, N$. However, the reciprocal is not true. Indeed, let

$$p(s) = \mathbf{j}s + \mathbf{k} \quad \text{and} \quad q(s) = \mathbf{j}s + \mathbf{j}$$

be two quaternionic polynomials. It is obvious that $\mathbf{j} \sim \mathbf{j}$ and $\mathbf{k} \sim \mathbf{j}$. Suppose that there exists a nonzero quaternion α such that $p(s) = \alpha q(s)\alpha^{-1}$, i.e., $p(s)\alpha = \alpha q(s)$. This would imply that $\alpha\mathbf{j} = \mathbf{j}\alpha$ and $\alpha\mathbf{k} = \mathbf{j}\alpha$. Then $\alpha\mathbf{j} = \alpha\mathbf{k}$, i.e., $\alpha(\mathbf{j} - \mathbf{k}) = 0$ and therefore $\alpha = 0$, which is impossible. $\square$

In the next definition we introduce the concept of total divisor in terms of equivalence classes.

Definition 2.2.17. The polynomial $d \in \mathbb{H}[s]$ is a *total divisor* of $p \in \mathbb{H}[s]$, $d \,\|\, p$, and p is a *total multiple* of d, if $[d] \,|\, [p]$, i.e., if for any $d' \in [d]$ and $p' \in [p]$, $d' \,|\, p'$.

Remark 2.2.18. It is obvious that a quaternionic polynomial is a divisor of itself. However, in general, this fact does not hold for total divisors, i.e., a quaternionic polynomial may not be a total divisor of itself. For instance, $(s + \mathbf{i}) \not\parallel (s + \mathbf{i})$. Indeed, the polynomial $s - \mathbf{i}$ is similar to $s + \mathbf{i}$ since $s - \mathbf{i} = \mathbf{j}(s + \mathbf{i})\mathbf{j}^{-1}$ but $s - \mathbf{i} = (s + \mathbf{i}) - 2\mathbf{i}$, i.e., $(s + \mathbf{i}) \nmid (s - \mathbf{i})$. We shall later prove, in Remark 2.2.28, that for every monic $p \in \mathbb{H}[s]$, $p \,\|\, p$ if and only if $p \in \mathbb{R}[s]$. $\square$

We next show that the definition of total divisor is equivalent to similar but simpler conditions.

Lemma 2.2.19. *Let $p, q \in \mathbb{H}[s]$. Then, the following are equivalent.*

$$1)\ p \,\|\, q; \qquad 2)\ p \,|\, [q]; \qquad 3)\ [p] \,|\, q.$$

Proof. Obviously the total divisor condition is sufficient for the other two. We prove that it is also necessary.

Suppose that $p \,|\, [q]$ and let $p' \in [p]$ and $q' \in [q]$. We shall prove that $p' \,|\, q'$.

By the definition of $[p]$ we know that there exists $\alpha \in \mathbb{H}$ such that $p' = \alpha p \alpha^{-1}$. Moreover, since $\alpha^{-1} q' \alpha \in [q]$, it follows from the hypothesis that there exists $d \in \mathbb{H}[s]$ such that $\alpha^{-1} q' \alpha = pd$. Therefore, letting $d' = \alpha d \alpha^{-1}$, we get that

$$q' = \alpha p \alpha^{-1} \alpha d \alpha^{-1} = p' d',$$

i.e., $p' \,|_l\, q'$. The proof that $p' \,|_r\, q'$ is similar, thus $p \,\|\, q$.

The same kind of reasoning also allows to show that $[p] \,|\, q$ implies that $p \,\|\, q$. □

Each quaternionic polynomial can be associated with two real polynomials that play an important role in the next chapter.

Definition 2.2.20. Given $p \in \mathbb{H}[s]$, we denote by $\mathcal{F}_p \in \mathbb{R}[s]$ the *greatest monic real factor* of p and by $\mathcal{M}_p \in \mathbb{R}[s]$ the *least monic real multiple* of p. We also denote by $\mathcal{Q}_p \in \mathbb{H}[s]$ the *quaternionic polynomial factor* of p such that $p = \mathcal{F}_p \mathcal{Q}_p$.

The relation between the aforementioned polynomials is given in the following result.

Lemma 2.2.21. *Let $p \in \mathbb{H}[s]$ be monic. Then*

1. $\mathcal{M}_p = p\overline{\mathcal{Q}_p} = \mathcal{F}_p \mathcal{Q}_p \overline{\mathcal{Q}_p}$.

2. $p\overline{p} = \mathcal{F}_p \mathcal{M}_p$.

Proof. 1. Note first that since p is monic so is $\mathcal{Q}_p$. It is clear, by Proposition 2.1.3-2, that $\mathcal{F}_p \mathcal{Q}_p \overline{\mathcal{Q}_p}$ is a monic real multiple of p. We just need to show that it is the least one. Suppose that $M \in \mathbb{R}[s]$ is another monic real multiple of p, i.e., $M = ap = a\mathcal{F}_p\mathcal{Q}_p$ for some $a \in \mathbb{H}[s]$. We claim that $\mathcal{F}_p\mathcal{Q}_p\overline{\mathcal{Q}_p} \,|\, M$. In order to show this it is enough to prove that every divisor of degree 1 of $\mathcal{F}_p\mathcal{Q}_p\overline{\mathcal{Q}_p}$ is a divisor of M. Let then $d(s) = s-\alpha$, $\alpha \in \mathbb{H}$, be a divisor of $\mathcal{F}_p\mathcal{Q}_p\overline{\mathcal{Q}_p}$. If d divides $\mathcal{F}_p$, clearly $d \,|\, M$ because $\mathcal{F}_p \,|\, p$ and $p \,|\, M$.

On the other hand, let d be a divisor of $\mathcal{Q}_p\overline{\mathcal{Q}_p}$, or equivalently of $\overline{\mathcal{Q}_p}\mathcal{Q}_p$. By Proposition 2.2.3, $\big(\overline{\mathcal{Q}_p}\mathcal{Q}_p\big)(\alpha) = 0$ and thus, since $\overline{\mathcal{Q}_p}\mathcal{Q}_p \in \mathbb{R}[s]$, by Lemma 2.2.10-4 this implies that $\psi_{[\alpha]} \,|\, \overline{\mathcal{Q}_p}\mathcal{Q}_p$. Then, by Lemma 2.2.10-5 there exists $\alpha' \in [\alpha]$ such that $\mathcal{Q}_p(\alpha') = 0$ and, as a consequence of Proposition 2.2.3, $M(\alpha') = 0$ since $\mathcal{Q}_p$ is a right divisor of M. Therefore, by Lemma 2.2.10-4 we have that $\psi_{[\alpha]} = \psi_{[\alpha']} \,|\, M$, i.e., $(s-\overline{\alpha})(s-\alpha) \,|\, M$ which implies that $d \,|\, M$.

2. Since $\overline{p} = \overline{\mathcal{F}_p\mathcal{Q}_p} = \mathcal{F}_p\overline{\mathcal{Q}_p}$ the result follows from 1.

$\square$

Example 2.2.22. Let $p = s^3+2\mathbf{j}s^2+s+2\mathbf{j} \in \mathbb{H}[s]$. The polynomial p can be factorized as $p = (s^2+1)(s+2\mathbf{j})$ and therefore it is easy to conclude that

$$\mathcal{F}_p = s^2+1 \quad \text{and} \quad \mathcal{M}_p = (s^2+1)(s+2\mathbf{j})(s-2\mathbf{j}) = (s^2+1)(s^2+4).$$

$\square$

In Proposition 1.1.3 it is stated that for every $\eta \in \mathbb{H}$ there exists a $z \in \mathbb{C}$ such that $z \in [\eta]$, i.e., such that $\operatorname{Re} z = \operatorname{Re}\eta$ and $|z| = |\eta|$. The following result can be regarded as a generalization of this property to polynomials.

Lemma 2.2.23. *For any monic quaternionic polynomial $q \in \mathbb{H}[s]$ there always exists a complex polynomial $p \in \mathbb{C}[s]$ such that $\mathcal{F}_p = \mathcal{F}_q$ and $\mathcal{M}_p = \mathcal{M}_q$. Moreover, if $p \in \mathbb{R}[s]$, then also $q \in \mathbb{R}[s]$ and $q = p$.*

Proof. Let $q = \mathcal{F}_q\mathcal{Q}_q \in \mathbb{H}[s]$. Note that, by the definition of $\mathcal{F}_q$, the polynomial $\mathcal{Q}_q$ has no real zeros and this clearly implies that also $\mathcal{Q}_q\overline{\mathcal{Q}_q}$ has no real zeros. Indeed, if $\big(\mathcal{Q}_q\overline{\mathcal{Q}_q}\big)(\alpha) = 0$ for some $\alpha \in \mathbb{R}$ then, since $s-\alpha \in \mathbb{R}[s]$ we may conclude by

Proposition 2.2.3 that $(s-\alpha) \mid \mathcal{Q}_q\overline{\mathcal{Q}_q}$. But, by definition, $\psi_\alpha = s-\alpha$ and hence, by Lemma 2.2.10-5 and from the fact that $[\alpha]=\{\alpha\}$ we have that $\mathcal{Q}_q(\alpha)=0$ leading thus to an absurd.

Since $\mathcal{Q}_q\overline{\mathcal{Q}_q} \in \mathbb{R}[s]$ and has no real zeros, there exists a complex polynomial $d \in \mathbb{C}[s]$ with no real zeros such that $\mathcal{Q}_q\overline{\mathcal{Q}_q} = d\overline{d}$. Consider $p = \mathcal{F}_q d \in \mathbb{C}[s]$. By the definition of $\mathcal{F}_q$ and d we have that $\mathcal{F}_p = \mathcal{F}_q$ and $\mathcal{Q}_p = d$. Moreover, by Lemma 2.2.21,

$$\mathcal{M}_q = \mathcal{F}_q\mathcal{Q}_q\overline{\mathcal{Q}_q} = \mathcal{F}_p\mathcal{Q}_p\overline{\mathcal{Q}_p} = \mathcal{M}_p.$$

If $p \in \mathbb{R}[s]$, then $\mathcal{Q}_p = \mathcal{Q}_q = 1$ and the result follows. $\square$

Lemma 2.2.24. *If $p \in \mathbb{C}[s]$, then $\mathcal{Q}_p$ and $\overline{\mathcal{Q}_p}$ are coprime.*

Proof. Note that since $p \in \mathbb{C}[s]$, so do $\mathcal{Q}_p$ and $\overline{\mathcal{Q}_p}$ and therefore there is no distinction between right and left coprimeness for these two polynomials.

Let $p \in \mathbb{C}[s]$ and suppose that $\mathcal{Q}_p$ and $\overline{\mathcal{Q}_p}$ are not coprime, i.e., there exists $\alpha \in \mathbb{C}\backslash\mathbb{R}$ such that $(s-\alpha) \mid \mathcal{Q}_p$ and $(s-\alpha) \mid \overline{\mathcal{Q}_p}$. Note that $\alpha \notin \mathbb{R}$ because $\mathcal{Q}_p$ has no real zeros. Then there exists $d \in \mathbb{C}[s]$ such that $\mathcal{Q}_p = d(s)(s-\alpha)$ which implies that $\overline{\mathcal{Q}_p} = (s-\overline{\alpha})\overline{d}(s)$. But $(s-\alpha) \mid \overline{\mathcal{Q}_p}$, and therefore $\overline{d}(s) = d''(s)(s-\alpha)$ for some $d'' \in \mathbb{C}[s]$. Hence,

$$\overline{\mathcal{Q}_p} = (s-\overline{\alpha})d''(s)(s-\alpha) = d''(s)(s-\overline{\alpha})(s-\alpha) = d''(s)\psi_{[\alpha]},$$

which, since $\psi_{[\alpha]} \in \mathbb{R}[s]$, contradicts the definition of $\mathcal{Q}_p$. $\square$

If $p \in \mathbb{H}[s] \setminus \mathbb{C}[s]$, the conclusion of the previous lemma does not necessarily hold as can be seen in the following example.

Example 2.2.25. Let

$$p = (s+\mathbf{i})(s-\mathbf{j})(s-\mathbf{i}) = s^3 - \mathbf{j}s^2 + (1-2\mathbf{k})s + \mathbf{j} \in \mathbb{H}[s].$$

First we show that $\mathcal{Q}_p = p$. Since $(s-\mathbf{i}) \mid_r p$, $p(\mathbf{i}) = 0$. Moreover, $\psi_{[\mathbf{i}]} = s^2+1 \not\mid p$ since

$$p = (s^2+1)(s+\mathbf{j}) - 2\mathbf{k}s.$$

Thus, by Lemma 2.2.10-2, $\mathbf{i}$ is the unique zero of p within its equivalence class. But, by Proposition 2.2.6, we have that all the zeros of p are equivalent to $\mathbf{i}$. Therefore $\mathbf{i}$ is the unique zero of p. This implies that p has not a real factor, i.e., $\mathcal{F}_p = 1$ and hence $\mathcal{Q}_p = p$. Thus $\overline{\mathcal{Q}_p} = \overline{p} = (s+\mathbf{i})(s+\mathbf{j})(s-\mathbf{i})$ and $\mathcal{Q}_p$ and $\overline{\mathcal{Q}_p}$ are not right coprime, since they share the right factor $(s-\mathbf{i})$. $\square$

The concept of total divisor has been introduced long ago in a different way. In fact, the last three conditions of the next theorem are, respectively, the definitions of total divisor given by Teichmüller [48], by Jacobson [23] and by Cohn [8]. In this theorem the equivalence between our total division condition and the other three is proven. The equivalence between the second and the third conditions was already presented in [23], but, in our opinion, with a wrong proof. First we introduce the notion of two-sided ideal.

Definition 2.2.26. If $\mathcal{I}$ is a subring of $\mathbb{H}[s]$ and $\mathbb{H}[s]\mathcal{I} \subseteq \mathcal{I}$ $(\mathcal{I}\mathbb{H}[s] \subseteq \mathcal{I})$ then $\mathcal{I}$ is called a *left (right) ideal* of $\mathbb{H}[s]$. If $\mathcal{I}$ is both a left and a right ideal, then $\mathcal{I}$ is said to be a *two-sided ideal.*

Theorem 2.2.27. *Let $p, q \in \mathbb{H}[s]$. Then the following conditions are equivalent:*

(i) $p \,\|\, q$;

(ii) $\mathbb{H}[s]q\mathbb{H}[s] \subseteq p\mathbb{H}[s] \cap \mathbb{H}[s]p$;

(iii) $\mathbb{H}[s]q \subseteq \mathcal{I} \subseteq \mathbb{H}[s]p$ *for some two-sided ideal* $\mathcal{I}$;

(iv) $q = abp$ *with* $bp \in \mathbb{R}[s]$ *and* $a, b \in \mathbb{H}[s]$.

Proof. We will show that the implications $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$ hold true.

$(i) \Rightarrow (ii)$ Assume that condition (i) is satisfied. We first prove that

$$\mathbb{H}q\mathbb{H} \subseteq p\mathbb{H}[s] \cap \mathbb{H}[s]p. \tag{2.4}$$

Indeed, by (i) and Lemma 2.2.19, $p \mid [q]$ and so, for any nonzero $\alpha \in \mathbb{H}$, there exists $d \in \mathbb{H}[s]$ such that $\alpha q \alpha^{-1} = pd$. Therefore, for any $\beta \in \mathbb{H}$,

$$\alpha q \beta = \alpha q \alpha^{-1} \alpha \beta = pd\alpha\beta \in p\mathbb{H}[s].$$

Analogously, we can prove that $\alpha q \beta \in \mathbb{H}[s]p$. Thus we only need to prove that (2.4) implies condition (ii).

Actually, for any $a = \sum \alpha_n s^n$, $b = \sum \beta_m s^m \in \mathbb{H}[s]$, equation (2.4) implies that there exist polynomials $l_{nm}, r_{nm} \in \mathbb{H}[s]$ such that

$$\alpha_n q \beta_m = l_{nm} p = p r_{nm}.$$

Therefore, recalling that $\eta s = s \eta$, $\eta \in \mathbb{H}$,

$$aqb = \sum \alpha_n q \beta_m s^{n+m} = \sum l_{nm} s^{n+m} p = p \sum r_{nm} s^{n+m},$$

showing that (ii) holds.

$(ii) \Rightarrow (iii)$ The condition is satisfied with $\mathcal{I}$ being the smallest ideal containing $\mathbb{H}[s]q\mathbb{H}[s]$, which can be shown to be contained in $p\mathbb{H}[s] \cap \mathbb{H}[s]p$.

$(iii) \Rightarrow (iv)$ We first show that the monic left and the right generators of any two-sided ideal $\mathcal{I}$ of $\mathbb{H}[s]$ are the same. Suppose that $\mathcal{I} = \mathbb{H}[s]g = g'\mathbb{H}[s]$. Then $g = g'h'$ and $g' = hg$ for some $h, h' \in \mathbb{H}[s]$. Thus, $g = hgh'$ which, taking into account that the degrees of g and g' must be equal, implies that h and h' are constant. Since g and g' are monic, $h = h' = 1$ and therefore $g = g'$.

Next we show that $g \in \mathbb{R}[s]$. Let $g = \mathcal{F}_g \mathcal{Q}_g$, where $\mathcal{F}_g \in \mathbb{R}[s]$ and $\mathcal{Q}_g \in \mathbb{H}[s]$. Suppose that $g \notin \mathbb{R}[s]$ and can hence be factorized as $\mathcal{Q}_g(s) = \mathcal{Q}'_g(s)(s - \alpha)$, for some $\alpha \in \mathbb{H} \backslash \mathbb{R}$.

Note that $\mathcal{Q}_g(\beta) \neq 0$ for every $\beta \in [\alpha]$ such that $\beta \neq \alpha$. Actually, by Lemma 2.2.10-3, if $\mathcal{Q}_g(\beta) = 0$ for some $\beta \in [\alpha] \backslash \{\alpha\}$, the minimal polynomial of $[\alpha]$, $\psi_{[\alpha]}$, would divide $\mathcal{Q}_g$, which is impossible by the definition of $\mathcal{F}_g$.

Let now $\alpha' \in [\alpha]$ be such that $\alpha' \neq \alpha$ and $\alpha' \neq \overline{\alpha}$ and consider the quaternionic polynomial $g(s)(s-\alpha') \in \mathcal{I}$. Since $\mathcal{I}$ is a two-sided ideal, there must exist $x \in \mathbb{H}[s]$ such that

$$x(s)g(s) = g(s)(s-\alpha') = \mathcal{F}_g(s)\mathcal{Q}_g(s)(s-\alpha'). \tag{2.5}$$

On the other hand

$$x(s)g(s) = x(s)\mathcal{F}_g(s)\mathcal{Q}_g(s) = \mathcal{F}_g(s)x(s)\mathcal{Q}_g(s), \tag{2.6}$$

From (2.5) and (2.6), we conclude that

$$x(s)\mathcal{Q}_g(s) = \mathcal{Q}_g(s)(s-\alpha'). \tag{2.7}$$

Since by hypothesis α is a zero of $\mathcal{Q}_g(s)$, and in turn of $x(s)\mathcal{Q}_g(s)$, a contradiction is achieved if we prove that α cannot be a zero of $\mathcal{Q}_g(s)(s-\alpha')$. Indeed, if α were a zero of $\mathcal{Q}_g(s)(s-\alpha')$, by Lemma 2.2.10-3, $\psi_{[\alpha']} \mid \mathcal{Q}_g(s)(s-\alpha')$ since $\alpha' \in [\alpha]$ is also a zero of $\mathcal{Q}_g(s)(s-\alpha')$. Consequently $(s-\overline{\alpha'}) \mid \mathcal{Q}_g(s)$ and thus, by Proposition 2.2.3, $\mathcal{Q}_g(\overline{\alpha'}) = 0$. But, since $\mathcal{Q}_g(s)$ has no zeros different from α belonging to $[\alpha]$, we conclude that α cannot be a zero of $\mathcal{Q}_g(s)(s-\alpha')$, which contradicts (2.7). This means that $g(s)$ must be a real polynomial.

As $\mathcal{I} \subseteq \mathbb{H}[s]p$, we have that $g = bp \in \mathbb{R}[s]$, for some $b \in \mathbb{H}[s]$. Finally, since $q \in \mathcal{I}$, there exists $a \in \mathbb{H}[s]$ such that $q = ag = abp$.

(iv) $\Rightarrow$ *(i)* By Lemma 2.2.19 we just need to prove that $[p] \mid q$. Let $p' \in [p]$, i.e., $p' = \eta p \eta^{-1}$ for some nonzero $\eta \in \mathbb{H}$. By Proposition 2.1.3-2, the fact that $bp(s)$ is a real polynomial implies that $bp = pb \in \mathbb{R}[s]$ and thus

$$q = abp = apb = pba.$$

Moreover, if we put $d = \eta b \eta^{-1} a$, we get that

$$q = pba = \eta\eta^{-1}pba = \eta p b \eta^{-1} a = \eta p \eta^{-1} \eta b \eta^{-1} a = p'd,$$

which means that $[p] \mid_l q$. Similarly we can prove that $[p] \mid_r q$ and the result follows. $\square$

Remark 2.2.28. Applying the fourth condition of the previous theorem it is easy to prove that if $p \in \mathbb{H}[s]$ is monic, then $p \,\|\, p$ if and only if $p \in \mathbb{R}[s]$. Indeed, if $p \,\|\, p$ then by Theorem 2.2.27-(iv) we have that $p = abp$ with $bp \in \mathbb{R}[s]$ and $a, b \in \mathbb{H}[s]$. Thus $ab = 1$ and $b \in \mathbb{H} \setminus \{0\}$. Taking into account that $bp \in \mathbb{R}[s]$ and p is monic, this implies that $b \in \mathbb{R}$ and hence $p \in \mathbb{R}[s]$. The other implication is obvious. $\square$

In general it is not easy to check, by definition, wether a polynomial is or is not a total divisor of another one. However, when the polynomials are factorized, the fourth condition of the previous theorem immediately allows to conclude about that. This condition can be stated in a more compact way using the notions of greatest monic real factor and least monic real multiple introduced in Definition 2.2.20.

Proposition 2.2.29. *Let $p, q \in \mathbb{H}[s]$. Then the following conditions are equivalent.*

$$(i)\ p \,\|\, q; \qquad (ii)\ p \,|\, \mathcal{F}_q; \qquad (iii)\ \mathcal{M}_p \,|\, \mathcal{F}_q; \qquad (iv)\ \mathcal{M}_p \,|\, q.$$

Proof. We will show that the following implications hold true: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i). Assume without loss of generality that the polynomials p and q are monic.

(i) $\Rightarrow$ (ii) If $p \,\|\, q$, by Theorem 2.2.27-(iv), we have that $q = abp$, where $bp \in \mathbb{R}[s]$ and $a \in \mathbb{H}[s]$. Moreover, since $a = \mathcal{F}_a \mathcal{Q}_a = \mathcal{Q}_a \mathcal{F}_a$ with $\mathcal{Q}_a \in \mathbb{H}[s]$ and $\mathcal{F}_a \in \mathbb{R}[s]$, $q = \mathcal{Q}_a \mathcal{F}_a bp$ which implies that $\mathcal{F}_q = \mathcal{F}_a bp$, and hence that $p | \mathcal{F}_q$.

(ii) $\Rightarrow$ (iii) If $p \,|\, \mathcal{F}_q$, there exists $a \in \mathbb{H}[s]$ such that $\mathcal{F}_q = ap = a\mathcal{Q}_p \mathcal{F}_p$. Note that, $\mathcal{F}_q \in \mathbb{R}[s]$ and therefore $a\mathcal{Q}_p \in \mathbb{R}[s]$, which implies that $a = a'\overline{\mathcal{Q}_p}$, for some $a' \in \mathbb{R}[s]$. Hence, by Lemma 2.2.21,

$$\mathcal{F}_q = a'\overline{\mathcal{Q}_p}\mathcal{Q}_p \mathcal{F}_p = a'\mathcal{M}_p,$$

showing that $\mathcal{M}_p \,|\, \mathcal{F}_q$.

(iii) $\Rightarrow$ (iv) This implication is obvious since $\mathcal{F}_q \,|\, q$.

$(iv) \Rightarrow (i)$ Assume that $\mathcal{M}_p \,|\, q$, i.e., that there exists $a \in \mathbb{H}[s]$ such that $q = a\mathcal{M}_p$. Since by Lemma 2.2.21 $\mathcal{M}_p = \overline{\mathcal{Q}_p}p$, $q = a\overline{\mathcal{Q}_p}p$ and, taking into account that $\overline{\mathcal{Q}_p}p \in \mathbb{R}[s]$, by Theorem 2.2.27-$(iv)$, we conclude that $p \,\|\, q$. □

Further properties of the polynomials $\mathcal{F}_p$, $\mathcal{M}_p$, and $\mathcal{Q}_p$ are stated in the following propositions.

Proposition 2.2.30. *Given two monic quaternionic polynomials $p, q \in \mathbb{H}[s]$,*

1. *The following three conditions are equivalent*
   - (i) $\mathcal{F}_{pq} = \mathcal{F}_p\mathcal{F}_q$;
   - (ii) $\mathcal{M}_{pq} = \mathcal{M}_p\mathcal{M}_q$;
   - (iii) $\left(\overline{\mathcal{Q}_p}, \mathcal{Q}_q\right)$ *are left coprime.*
2. $pq \in \mathbb{R}[s]$ *if and only if* $\overline{\mathcal{Q}_p} = \mathcal{Q}_q$.

Proof. 1. First note that $pq = \mathcal{F}_p\mathcal{Q}_p\mathcal{F}_q\mathcal{Q}_q = \mathcal{F}_p\mathcal{F}_q\mathcal{Q}_p\mathcal{Q}_q$, hence

$$\mathcal{F}_{pq} = \mathcal{F}_p\mathcal{F}_q\mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q}$$

and therefore

$$\mathcal{F}_{pq} = \mathcal{F}_p\mathcal{F}_q \Leftrightarrow \mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q} = 1.$$

Analogously we can prove that

$$\mathcal{M}_{pq} = \mathcal{M}_p\mathcal{M}_q \Leftrightarrow \mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q} = 1,$$

yielding that $(i) \Leftrightarrow (ii)$. In order to prove that (iii) is equivalent to (i) and (ii), we only have to show that

$$\mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q} = 1 \Leftrightarrow \left(\overline{\mathcal{Q}_p}, \mathcal{Q}_q\right) \text{ are left coprime} \tag{2.8}$$

If $\overline{\mathcal{Q}_p}$ and $\mathcal{Q}_q$ have a non-trivial left common factor, say x, then $\overline{x}$ is a right factor of $\mathcal{Q}_p$ and $\mathcal{Q}_p\mathcal{Q}_q$ is a multiple of the real polynomial $\overline{x}x$. Thus $\overline{x}x \,|\, \mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q}$, which proves the direct implication.

To prove the converse implication, assume that $\mathcal{F}_{\mathcal{Q}_p\mathcal{Q}_q} \neq 1$. Then, there exists $\lambda \in \mathbb{H}$ such that $\psi_{[\lambda]}|\mathcal{Q}_p\mathcal{Q}_q$. Note that $\lambda \notin \mathbb{R}$ because otherwise $\psi_{[\lambda]} = s-\lambda$ and $(s-\lambda) \mid \mathcal{Q}_p$ or $(s-\lambda) \mid \mathcal{Q}_q$, which is impossible by the definition of $\mathcal{Q}_p$ and $\mathcal{Q}_q$. Moreover, if $\mathcal{Q}_p(\nu) \neq 0$ for every $\nu \in [\lambda]$ then by Lemma 2.2.10-2, $\psi_{[\lambda]}$ is not a divisor of $\mathcal{Q}_p$, and must be a divisor of $\mathcal{Q}_q$ (since it is a real divisor of $\mathcal{Q}_p\mathcal{Q}_q$). But this is impossible since $\psi_{[\lambda]} \in \mathbb{R}[s]$ and $\mathcal{Q}_q$ does not allow real factors. Therefore, $\mathcal{Q}_p$ must have a zero $\nu \in [\lambda]$, i.e., $\mathcal{Q}_p = ax$ with $x(s) = s - \nu$ and $a \in \mathbb{H}[s]$. Note that $x\overline{x} = \psi_{[\nu]} = \psi_{[\lambda]}$.

By the division algorithm, there exist $y \in \mathbb{H}[s]$ and $\eta \in \mathbb{H}$ such that $\overline{\mathcal{Q}_q} = yx + \eta$, thus

$$\mathcal{Q}_p\mathcal{Q}_q = ax\overline{x}\,\overline{y} + \mathcal{Q}_p\overline{\eta} = a\overline{y}\psi_{[\lambda]} + \mathcal{Q}_p\overline{\eta}.$$

So, as $\psi_{[\lambda]}|\mathcal{Q}_p\mathcal{Q}_q$, we have that $\psi_{[\lambda]}|\mathcal{Q}_p\overline{\eta}$, which is possible if and only if $\eta = 0$. Therefore, x is a common right factor of $\mathcal{Q}_p$ and $\overline{\mathcal{Q}_q}$, i.e., $\overline{x}$ is a common left factor of $\overline{\mathcal{Q}_p}$ and $\mathcal{Q}_q$.

2. The "if" statement is trivial. In order to prove the "only if" part assume that $pq \in \mathbb{R}[s]$ and let $x \in \mathbb{H}[s]$ be the greatest monic left common factor of $\overline{\mathcal{Q}_p}$ and $\mathcal{Q}_q$, i.e., $\overline{\mathcal{Q}_p} = xa$ and $\mathcal{Q}_q = xb$ where $(a, b) \in \mathbb{H}[s] \times \mathbb{H}[s]$ are left coprime. Then,

$$pq = \mathcal{F}_p\mathcal{Q}_p\mathcal{F}_q\mathcal{Q}_q = \mathcal{F}_p\overline{a}\,\overline{x}\mathcal{F}_q xb = \mathcal{F}_p\mathcal{F}_q\overline{a}\,\overline{x}xb = \mathcal{F}_p\mathcal{F}_q\overline{x}x\overline{a}b. \tag{2.9}$$

Since (a, b) are left coprime and $(\mathcal{Q}_a, \mathcal{Q}_b) = (a, b) = \left(\overline{\overline{a}}, b\right)$, by (2.8) it follows that $\mathcal{F}_{\overline{a}b} = 1$. By (2.9), $pq = \mathcal{F}_p\mathcal{F}_q\overline{x}x\overline{a}b$ and so, since $\mathcal{F}_p\mathcal{F}_q\overline{x}x \in \mathbb{R}[s]$,

$$\mathcal{F}_{pq} = \mathcal{F}_p\mathcal{F}_q\overline{x}x\mathcal{F}_{\overline{a}b} = \mathcal{F}_p\mathcal{F}_q\overline{x}x.$$

Moreover, $pq \in \mathbb{R}[s]$, i.e., $pq = \mathcal{F}_{pq}$ and then

$$pq = \mathcal{F}_{pq} = \mathcal{F}_p\mathcal{F}_q\overline{x}x. \tag{2.10}$$

From equations (2.9) and (2.10) it follows that $a = b = 1$, and hence $\overline{\mathcal{Q}_p} = \mathcal{Q}_q$. $\square$

Corollary 2.2.31. *If p and q are monic quaternionic polynomials such that $pq \in \mathbb{R}[s]$, then $pq = qp = \mathcal{F}_p\mathcal{M}_q = \mathcal{M}_p\mathcal{F}_q$.*

Proof. Assume that $pq \in \mathbb{R}[s]$. By Proposition 2.2.30-2, $\mathcal{Q}_q = \overline{\mathcal{Q}_p}$, thus $pq = p\mathcal{Q}_q\mathcal{F}_q = p\overline{\mathcal{Q}_p}\mathcal{F}_q$ and hence by Lemma 2.2.21-1 we have that $pq = \mathcal{M}_p\mathcal{F}_q$. Analogously we can prove that $qp = \mathcal{F}_p\mathcal{M}_q$. Finally, by Proposition 2.1.3-2, $pq = qp$, which concludes the proof. $\square$

Let us now define the multiplicity of the equivalence class $[\lambda]$ with respect to p as

$$\mu_{[\lambda]}(p) = \max\{\mu_\nu(p) : \nu \in [\lambda]\}.$$

This notion of multiplicity will play an important role in the study of the stability of a dynamical system that will be given in Chapter 5. It turns out that the multiplicity of an equivalence class with respect to $p \in \mathbb{H}[s]$ can be related to the multiplicity of the elements of this class as zeros of $\mathcal{M}_p$.

Proposition 2.2.32. *For any $p \in \mathbb{H}[s]$ and $\nu \in [\lambda]$, $\mu_{[\lambda]}(p) = \mu_\nu(\mathcal{M}_p)$.*

Proof. The fact is trivial if $\lambda \in \mathbb{R}$. So, let $\lambda \notin \mathbb{R}$ and suppose, without loss of generality, that p is monic, further let $n = \mu_{[\lambda]}(p)$, i.e., $n = \mu_\alpha(p)$ for some $\alpha \in [\lambda]$ such that $\mu_\alpha(p) \geq \mu_\beta(p)$ for $\beta \in [\lambda]$. According to Remark 2.2.14, p has a maximum degree right factor $d \in \mathbb{H}[s]$ with unique zero α such that $d\overline{d} = \psi_{[\alpha]}^n$. Thus $p = qd$, for some $q \in \mathbb{H}[s]$. On the other hand, by Definition 2.2.20, $p = \mathcal{F}_p\mathcal{Q}_p$ and $\mathcal{Q}_p$ has no real factors. Suppose that $\mathcal{Q}_p = by$ where $b \in \mathbb{H}[s]$ and $y \in \mathbb{H}[s]$ is the maximal right factor of $\mathcal{Q}_p$ with unique zero α. If $\deg y = m$, by Lemma 2.2.11,

$$y(s) = (s - \alpha_m) \cdots (s - \alpha_2)(s - \alpha_1),\ \alpha_l \in [\alpha_1],\ \alpha_l \neq \overline{\alpha_{l-1}},\ l = 2, \dots, n,$$

with $\alpha_1 = \alpha$. Note that b cannot have zeros in $[\alpha]$. Indeed, let $b(\alpha') = 0$, for some $\alpha' \in [\alpha]$, i.e., $b(s) = b'(s)(s - \alpha')$. If $\alpha' = \overline{\alpha_m}$, then

$$\mathcal{Q}_p(s) = b(s)y(s) = b'(s)(s - \overline{\alpha_m})(s - \alpha_m) \cdots (s - \alpha_2)(s - \alpha),$$

and therefore $\psi_{[\alpha_m]} = (s - \overline{\alpha_m})(s - \alpha_m) \in \mathbb{R}[s]$ is a real factor of $\mathcal{Q}_p$, leading to an absurd. If $\alpha' \neq \overline{\alpha_m}$, by Lemma 2.2.11, the polynomial $(s - \alpha)y(s)$ is a right factor of

$\mathcal{Q}_p$ with unique zero α and has degree $m+1$, impossible by the definition of y. Hence b cannot have zeros in $[\alpha]$. Since $y \,|_r\, \mathcal{Q}_p \,|_r\, p$, y is a right divisor of p with unique zero α and therefore, by the definition of d, $y \,|_r\, d$. Thus $d = xy$ for some $x \in \mathbb{H}[s]$ whose only factors are $s - \alpha_l$, $\alpha_l \in [\alpha]$. So,

$$p = \mathcal{F}_p \mathcal{Q}_p = \mathcal{F}_p b y \qquad \text{and} \qquad p = qd = qxy$$

which implies that $\mathcal{F}_p b = qx$ and thus $x \,|_r\, \mathcal{F}_p b$. Since the polynomials x and b cannot have common factors, $x \,|_r\, \mathcal{F}_p$. Moreover, x does not have real factors and, by Lemma 2.2.21-2, this implies that $\mathcal{M}_x = x\overline{x}$. Therefore, by Proposition 2.2.29, $x\overline{x} \,|_r\, \mathcal{F}_p$. So, $\mathcal{F}_p = a x \overline{x} = a \overline{x} x$ for some $a \in \mathbb{R}[s]$ and then

$$p = \mathcal{F}_p \mathcal{Q}_p = a \overline{x} x b y = a b \overline{x} x y = a b \overline{x} d.$$

We prove next that also a cannot have zeros in $[\alpha]$. Suppose that $a(\alpha') = 0$, for some $\alpha' \in [\alpha]$. Since $a \in \mathbb{R}[s]$, by Lemma 2.2.10-4, $\psi_{[\alpha]} = \psi_{[\alpha']} \,|\, a$, which implies that $a = a' \psi_{[\alpha]}$, $a' \in \mathbb{R}[s]$. Then

$$p = a b \overline{x} d = a' \psi_{[\alpha]} b \overline{x} d = a' b \overline{x} \psi_{[\alpha]} d.$$

But, by Lemma 2.2.15-2, $\mu_\alpha(\psi_{[\alpha]} d) = \mu_\alpha(d) + 1 = n + 1$, which is impossible since $\mu_{[\alpha]}(p) = n$ and $\psi_{[\alpha]} d$ is a right factor of p. Thus a cannot have zeros in $[\alpha]$.

By Lemma 2.2.21,

$$\mathcal{M}_p = \mathcal{F}_p \mathcal{Q}_p \overline{\mathcal{Q}_p} = a x \overline{x} b y \overline{y}\, \overline{b} = a b \overline{b} x y \overline{y}\, \overline{x},$$

and then, since $\psi_{[\alpha]}^n = d\overline{d} = xy\overline{y}\,\overline{x}$,

$$\mathcal{M}_p = a b \overline{b} \psi_{[\alpha]}^n.$$

Let $\nu \in [\lambda] = [\alpha]$. Then $\psi_{[\alpha]} = \psi_{[\nu]}$ and since a and b have no zeros in $[\alpha]$,

$$\mu_\nu(\mathcal{M}_p) = \mu_\nu\left(a b \overline{b} \psi_{[\nu]}^n\right) = \mu_\nu\left(\psi_{[\nu]}^n\right).$$

Finally, by Lemma 2.2.15-1, $\mu_\nu\left(\psi_{[\nu]}^n\right) = n = \mu_{[\lambda]}(p)$. □

2.3 J-similarity

In the previous section a notion of similarity between quaternionic polynomials has been introduced. However another powerful similarity property has been considered in the literature. In order to distinguish it from the first one, we will call it J-similarity, and denote it by $\sim_J$, where the J stands for Jacobson, who first introduced this notion, [23].

This property is in particular important in the study of Smith forms for quaternionic polynomial matrices presented in the next chapter.

Definition 2.3.1. [8, 23] Two quaternionic polynomials $a, d \in \mathbb{H}[s]$ are said to be *J-similar*, $a \sim_J d$, if there exist $b, c \in \mathbb{H}[s]$ such that the relation

$$ab = cd$$

is a *coprime relation*. By this it is meant that (a, c) are left coprime and (b, d) are right coprime.

It is obvious that similarity implies J-similarity. Indeed, if two quaternionic polynomials $p \sim q$, i.e., $p = \alpha q \alpha^{-1}$, $\alpha \in \mathbb{H} \setminus \{0\}$, then $p\alpha = \alpha q$ with (p, α) left coprime and (α, q) right coprime which implies that $p \sim_J q$. However the reciprocal does not hold as shown in the following example.

Example 2.3.2. Consider the quaternionic polynomials

$$p(s) = \left(s + \frac{3\mathbf{i} + 4\mathbf{j}}{5}\right)(s - 2\mathbf{j}) = s^2 + \frac{3\mathbf{i} - 6\mathbf{j}}{5}s + \frac{8 - 6\mathbf{k}}{5} \tag{2.11}$$

and

$$q(s) = (s - \mathbf{i})(s - 2\mathbf{j}) = s^2 - (\mathbf{i} + 2\mathbf{j})s + 2\mathbf{k}. \tag{2.12}$$

Note that $\frac{8-6\mathbf{k}}{5} \not\sim 2\mathbf{k}$ since their real parts are different (see Proposition 1.1.2). Therefore, by (2.11), (2.12) and Remark 2.2.16 this implies that $p \not\sim q$.

On the other hand,

$$p\overline{p} = \overline{q}q = (s^2 + 1)(s^2 + 4)$$

and it is not difficult to check that $(p, \overline{q})$ are left coprime and that $(\overline{p}, q)$ are right coprime. Hence $p \sim_J q$. □

Our next result, that will be relevant for Section 3.2, relates the real polynomials $a\overline{a}$ and $d\overline{d}$ in case $a \sim_J d$.

Proposition 2.3.3. *Let* $a = \sum_{l=0}^{n} a_l s^l, d = \sum_{l=0}^{m} d_l s^l \in \mathbb{H}[s]$ *be such that* $|a_n| = |d_m|$.

If $a \sim_J d$, *i.e., there exists* $b, c \in \mathbb{H}[s]$ *such that*

$$ab = cd \tag{2.13}$$

is a coprime relation, then

$$a\overline{a} = d\overline{d} \quad \text{and} \quad b\overline{b} = c\overline{c}.$$

Proof. Suppose first that a and d are monic.

Let $\alpha_1 \in \mathbb{H}$ be such that $d(\alpha_1) = 0$. By Proposition 2.2.3, $d(s) = \widetilde{d}(s)(s - \alpha_1)$, i.e., $c(s)d(s) = c(s)\widetilde{d}(s)(s - \alpha_1)$, which implies that $(cd)(\alpha_1) = 0$. Then, by (2.13), $(ab)(\alpha_1) = 0$ but $b(\alpha_1) \neq 0$ because (b, d) are right coprime. Thus, by Proposition 2.2.5 there exists $\alpha_1' = b(\alpha_1)\alpha_1 b(\alpha_1)^{-1} \sim \alpha_1$ such that $a(\alpha_1') = 0$, i.e., $a(s) = \widetilde{a}(s)(s - \alpha_1')$ for some $\widetilde{a} \in \mathbb{H}[s]$. This implies that $a(s)b(s) = \widetilde{a}(s)(s - \alpha_1')b(s)$. Moreover, by Proposition 2.2.5

$$\big((s - \alpha_1')b(s)\big)(\alpha_1) = (\alpha_1' - \alpha_1')b(\alpha_1) = 0$$

and hence $(s - \alpha_1')b(s) = \widetilde{b}(s)(s - \alpha_1)$, for some $\widetilde{b} \in \mathbb{H}[s]$. Thus

$$\begin{aligned} a(s)b(s) = c(s)d(s) \quad &\Leftrightarrow \quad \widetilde{a}(s)\widetilde{b}(s)(s - \alpha_1) = c(s)\widetilde{d}(s)(s - \alpha_1) \\ &\Leftrightarrow \quad \widetilde{a}(s)\widetilde{b}(s) = c(s)\widetilde{d}(s). \end{aligned}$$

Note that both $\widetilde{a}(s)$ and $\widetilde{d}(s)$ are monic. Proceeding analogously as many times as necessary it is possible to cancel out all the factors of $d(s)$, i.e, if

$$d(s) = (s - \alpha_m) \cdots (s - \alpha_1), \;\; \alpha_l \in \mathbb{H}, \;\; l = 1, \ldots, m, \tag{2.14}$$

we obtain $\widehat{a}(s)\widehat{b}(s) = c(s)$, with

$$\begin{cases} a(s) = \widehat{a}(s)(s-\alpha'_m)\cdots(s-\alpha'_1),\ \alpha'_l \sim \alpha_l,\ l=1,\dots,m \\ (s-\alpha'_m)\cdots(s-\alpha'_1)b(s) = \widehat{b}(s)(s-\alpha_m)\cdots(s-\alpha_1) \end{cases}. \tag{2.15}$$

Since $a(s)$ is monic so is $\widehat{a}(s)$. If we prove that $\widehat{a}(s) = 1$, by (2.14) and the first equation of (2.15) and as a consequence of Lemma 2.2.10-1, we have that $a\overline{a} = d\overline{d}$ as desired.

Since (a, c) are left coprime by hypothesis, it is clear that also $(\widehat{a}, c)$ are left coprime, which implies that their conjugates $(\overline{\widehat{a}}, \overline{c})$ are right coprime. Moreover,

$$\widehat{a}(s)\widehat{b}(s) = c(s) \ \Leftrightarrow\ \overline{\widehat{b}}(s)\overline{\widehat{a}}(s) = \overline{c}(s).$$

In the same way as the factors of $d(s)$ were cancelled out, it is possible to cancel out all the factors of $\overline{c}$ and, letting $\overline{c}(s) = (s-\beta_r)\cdots(s-\beta_1),\ \beta_l \in \mathbb{H},\ l = 1,\dots,r$, we get $b'(s)a'(s) = 1$, with

$$\begin{cases} \overline{\widehat{b}}(s) = b'(s)(s-\beta'_r)\cdots(s-\beta'_1),\ \beta'_l \sim \beta_l,\ l=1,\dots,r \\ (s-\beta'_r)\cdots(s-\beta'_1)\overline{\widehat{a}}(s) = a'(s)(s-\beta_r)\cdots(s-\beta_1) \end{cases}. \tag{2.16}$$

By the second equation of (2.16), since $\widehat{a}(s)$ is monic so is $a'(s)$. Moreover, $b'(s)a'(s) = 1$ implies that $a'(s) = 1$ and hence also $\overline{\widehat{a}} = 1$, i.e., $\widehat{a} = 1$ and thus, as we stated above,

$$a\overline{a} = d\overline{d}. \tag{2.17}$$

Furthermore $ab = cd \ \Leftrightarrow\ \overline{b}\overline{a} = \overline{d}\overline{c}$ which, by (2.17), implies that

$$ab\overline{b}\overline{a} = cd\overline{d}\overline{c} \ \Leftrightarrow\ a\overline{a}b\overline{b} = d\overline{d}c\overline{c} \ \Leftrightarrow\ a\overline{a}(b\overline{b} - c\overline{c}) = 0 \ \Leftrightarrow\ b\overline{b} = c\overline{c}.$$

If a and d are not monic, define the quaternions

$$\mu = \frac{\overline{a_n}}{|a_n|^2} \quad \text{and} \quad \nu = \frac{\overline{d_m}}{|d_m|^2}.$$

Then, due to the assumption that $|a_n| = |d_m|$,

$$\mu\overline{\mu} = \frac{\overline{a_n}a_n}{|a_n|^4} = \frac{1}{|a_n|^2} = \frac{1}{|d_m|^2} = \nu\overline{\nu}.$$

Define also the polynomials

$$\widetilde{a}(s) = a(s)\mu, \quad \widetilde{c}(s) = c(s)\nu^{-1}, \quad \widetilde{b}(s) = \mu^{-1}b(s), \quad \widetilde{d}(s) = \nu d(s).$$

Since by hypothesis (a, c) are left coprime and (b, d) are right coprime, also $(\widetilde{a}, \widetilde{c})$ are left coprime and $(\widetilde{b}, \widetilde{d})$ are right coprime. Now

$$a(s)b(s) = c(s)d(s) \;\Leftrightarrow\; a(s)\mu\mu^{-1}b(s) = c(s)\nu^{-1}\nu d(s) \;\Leftrightarrow\; \widetilde{a}(s)\widetilde{b}(s) = \widetilde{c}(s)\widetilde{d}(s),$$

where $\widetilde{a}(s)$ and $\widetilde{d}(s)$ are monic and $\widetilde{a} \sim_J \widetilde{d}$, and the result follows then from the first part of the proof. □

Chapter 3

Quaternionic polynomial and rational matrices

In this chapter we start by giving some definitions and preliminary results on quaternionic polynomials and rational matrices. Most of them are simply an extension of the ones concerning real or complex polynomial matrices. However, as expected (cf Section 1.2.2), this does not happens with the notion of determinant. Since such a notion was missing for quaternionic polynomial matrices, following the spirit of Dieudonné's approach mentioned in Section 1.2.2, we propose, in Section 3.2, a new definition for this case. In Section 3.3, we define the complex adjoint of a polynomial matrix and show that quaternionic polynomial matrices share many algebraic properties with their complex adjoint. In Section 3.4, we present the quaternionic Smith form of quaternionic polynomial matrices, which unlike the commutative case is not unique. Its source of non uniqueness was already investigated in [17, 23, 34] and we give here alternative characterizations. We also characterize the complex Smith form of complex adjoint matrices and give its relation with the quaternionic Smith form. Finally, in Section 3.5, we introduce the quaternionic Smith-McMillan form of a quaternionic rational matrix and the results of the previous section are extended to rational matrices.

As usual, $\mathbb{H}^{g\times r}[s]$ and $\mathbb{H}^{g\times r}[s,s^{-1}]$ will denote, respectively, the set of the $g\times r$ matrices with entries in $\mathbb{H}[s]$ and $\mathbb{H}[s,s^{-1}]$. As we did for polynomials, we may also omit here the indeterminate s and write $R\in\mathbb{H}^{g\times r}[s]$ if no ambiguity arises.

For the sake of simplicity, results and definitions that trivially generalize to L-polynomial matrices are only stated for polynomial matrices. In case a fundamental difference occurs both cases are presented.

3.1 Definitions and preliminary results

Similar to what happens for the commutative case, it turns out that unimodular and full row rank matrices play an important role in the study of the algebraic properties of quaternionic polynomial matrices and of dynamical systems that will be carried out in this thesis. Therefore, based in [43, 49], we start this section by introducing these notions as well as other basic definitions concerning quaternionic polynomial matrices.

Definition 3.1.1. A quaternionic polynomial matrix $U\in\mathbb{H}^{g\times g}[s]$ is said to be *unimodular* if it admits an inverse in $\mathbb{H}^{g\times g}[s]$, i.e., if there exists another matrix $V\in\mathbb{H}^{g\times g}[s]$ such that $VU=UV=I$.

Note that, since for V and U square,

$$VU=I \;\Leftrightarrow\; UV=I,$$

the condition of the definition may be replaced by $VU=I$ (or $UV=I$).

Remark 3.1.2. Consider the matrix

$$U=\begin{bmatrix}1 & 0\\ 0 & s\end{bmatrix}.$$

It is clear that, as an element of $\mathbb{H}^{2\times 2}[s]$, U is not a unimodular matrix. However, regarding U as a L-polynomial matrix, U is in fact unimodular since

$$V = \begin{bmatrix} 1 & 0 \\ 0 & s^{-1} \end{bmatrix} \in \mathbb{H}^{2\times 2}[s, s^{-1}]$$

and $UV = I$. $\square$

As referred in Section 2.1, given two quaternionic polynomials $p(s)$ and $q(s)$ and their product $r(s) = p(s)q(s)$, in general $r(\lambda) \neq p(\lambda)q(\lambda)$, $\lambda \in \mathbb{H}$. Hence, the same happens for quaternionic matrices. In particular, if $U(s) \in \mathbb{H}^{g\times g}[s]$ is unimodular, then $U(\lambda)$ is not necessarily invertible for all $\lambda \in \mathbb{H}$. Note that, in case $\lambda \in \mathbb{R}$, the matrix $U(\lambda)$ is invertible.

Example 3.1.3. Let

$$U = \begin{bmatrix} -\mathbf{i}s + \mathbf{k} & \mathbf{j}s \\ -\mathbf{i} & \mathbf{j} \end{bmatrix}.$$

The matrix U is unimodular since

$$U^{-1} = \begin{bmatrix} -\mathbf{k} & \mathbf{k}s \\ 1 & -s - \mathbf{j} \end{bmatrix}.$$

However, $U\left(1 + \frac{1}{2}\mathbf{j}\right)$ is not invertible. Indeed,

$$U\left(1 + \tfrac{1}{2}\mathbf{j}\right) \begin{bmatrix} 1 \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} -\mathbf{i} + \frac{1}{2}\mathbf{k} & -\frac{1}{2} + \mathbf{j} \\ -\mathbf{i} & \mathbf{j} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$\square$

Definition 3.1.4. Two matrices $R, R' \in \mathbb{H}^{g\times r}[s]$ are said to be *equivalent* if there exist unimodular matrices $U \in \mathbb{H}^{g\times g}[s]$ and $V \in \mathbb{H}^{r\times r}[s]$ such that

$$R = UR'V.$$

Definition 3.1.5. $R \in \mathbb{H}^{g\times r}[s]$ is said to have *full row rank* (frr) if its rows are linearly independent on the left, i.e.,

$$\forall\, x \in \mathbb{H}^{1\times g}[s], \quad xR = 0 \Rightarrow x = 0.$$

R has *full column rank* (fcr) if R^T has full row rank. If R is a square full row rank matrix, then we simply say that R has *full rank.*

As happens for quaternionic polynomials, it is possible to define the concepts of divisor, left coprimeness, etc, for quaternionic polynomial matrices.

Definition 3.1.6. Let $A \in \mathbb{H}^{n\times m}[s]$. A matrix $D \in \mathbb{H}^{n\times n}[s]$ is a *left divisor* of A, and A is a *right multiple* of D, if there exists a matrix $C \in \mathbb{H}^{n\times m}[s]$ such that $A = DC$. A matrix $A \in \mathbb{H}^{n\times m}[s]$ is *left prime* if every left divisor of A is unimodular.

Definition 3.1.7. Let $A \in \mathbb{H}^{n\times m}[s]$ and $B \in \mathbb{H}^{n\times p}[s]$. A matrix $D \in \mathbb{H}^{n\times n}[s]$ is a *greatest common left divisor* (gcld) of A and B if

1. D is a left divisor of both A and B, and
2. D is a right multiple of every common left divisor of A and B.

Two matrices $A \in \mathbb{H}^{n\times m}[s]$ and $B \in \mathbb{H}^{n\times p}[s]$ are *left coprime* if every gcld of A and B is unimodular.

The definitions of right divisor, greatest common right divisor (gcrd), and right coprimeness are entirely analogous.

The notation introduced in Notation 1.2.2 is extended to the polynomial case, i.e., we denote by $SL(n, \mathbb{H}[s])$ the set of all $n \times n$ polynomial matrices that can be decomposed as a product of matrices of the types P_{lm} and $B_{lm}(\alpha)$, $\alpha \in \mathbb{H}[s]$. Note that every matrix in $SL(n, \mathbb{H}[s])$ is unimodular.

In the sequel some preliminary results on quaternionic polynomial matrices are stated. The following technical lemma is relevant for our definition of determinant of quaternionic polynomial matrices that will be given in Section 3.2.

Lemma 3.1.8. *Let*

$$R = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \in \mathbb{H}^{2\times 1}[s].$$

Then there exists $U \in SL(2, \mathbb{H}[s])$ *such that*

$$UR = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} t \\ 0 \end{bmatrix},$$

with t *a* gcrd *of* γ_1 *and* γ_2, *and*

$$u_{21} \sim_J g \quad \textit{and} \quad |\operatorname{lc} u_{21}| = |\operatorname{lc} g|,$$

where $g \in \mathbb{H}[s]$ *is such that* $\gamma_2 = gt$.

Proof. If either $\gamma_1 = 0$ or $\gamma_2 = 0$ the result is easy to check. Let then $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$. Assume first that $\deg \gamma_1 \geq \deg \gamma_2$. Then, by the Euclidian division algorithm for quaternionic polynomials, there exist $d_1, r_1 \in \mathbb{H}[s]$ such that

$$\gamma_1 = d_1\gamma_2 + r_1, \quad \text{with} \quad \deg r_1 < \deg \gamma_2 \quad \text{or} \quad r_1 = 0. \tag{3.1}$$

Thus we may write

$$U_1 \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \gamma_2 \\ r_1 \end{bmatrix}, \tag{3.2}$$

with $U_1 = \begin{bmatrix} 0 & 1 \\ 1 & -d_1 \end{bmatrix}$.

If $r_1 = 0$, $\gamma_1 = d_1\gamma_2$ and therefore γ_2 is a gcrd of γ_1 and γ_2. Taking $U = U_1$, $u_{21} = g = 1$ and the result follows. If r_1 is nonzero, there exist again $d_2, r_2 \in \mathbb{H}[s]$ such that

$$\gamma_2 = d_2 r_1 + r_2, \quad \text{with} \quad \deg r_2 < \deg r_1 \quad \text{or} \quad r_2 = 0, \tag{3.3}$$

and hence, letting $U_2 = \begin{bmatrix} 0 & 1 \\ 1 & -d_2 \end{bmatrix}$,

$$U_2 \begin{bmatrix} \gamma_2 \\ r_1 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}. \tag{3.4}$$

If $r_2 = 0$, combining (3.1) and (3.3) yields

$$\gamma_1 = (d_1 d_2 + 1) r_1 \quad \text{and} \quad \gamma_2 = d_2 r_1$$

and, since $d_1 d_2 + 1$ and d_2 are right coprime, this implies that r_1 is a gcrd of γ_1 and γ_2. Moreover, by (3.2) and (3.4), if $U = U_2 U_1$

$$U \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 1 & -d_1 \\ -d_2 & 1 + d_2 d_1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ 0 \end{bmatrix}.$$

Now, $u_{21} = -d_2$, $g = d_2$ and therefore $|\operatorname{lc} u_{21}| = |\operatorname{lc} g_1|$.

If r_2 is nonzero, by continuing this procedure we obtain, after a finite number l of steps,

$$U_l \begin{bmatrix} r_{l-2} \\ r_{l-1} \end{bmatrix} = \begin{bmatrix} r_{l-1} \\ 0 \end{bmatrix}$$

and

$$\begin{aligned} \gamma_1 &= \widetilde{g}\, r_{l-1} \\ \gamma_2 &= g r_{l-1}, \quad r_0 = \gamma_2, \end{aligned} \tag{3.5}$$

with $\widetilde{g} = d_1 d_2 \cdots d_l +$ *lower order terms (l.o.t)* and $g = d_2 d_3 \cdots d_l + l.o.t.$

Thus

$$U \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} t \\ 0 \end{bmatrix}, \tag{3.6}$$

with $U = U_l \cdots U_1 = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \in SL(2, \mathbb{H}[s])$ and $t = r_{l-1}$. Further, it is not difficult to see that

$$\begin{aligned} u_{21} &= (-1)^{l-1} d_l \cdots d_2 + l.o.t \\ u_{22} &= (-1)^{l} d_l \cdots d_2 d_1 + l.o.t. \end{aligned} \tag{3.7}$$

Note that, by (3.5) and (3.6) we can conclude that t is a gcrd of γ_1 and γ_2. Indeed, let $x \in \mathbb{H}[s]$ be any common right divisor of γ_1 and γ_2, i.e., $\gamma_1 = y_1 x$ and $\gamma_2 = y_2 x$ for some $y_1, y_2 \in \mathbb{H}[s]$. By (3.6) we have that $u_{11}\gamma_1 + u_{12}\gamma_2 = t$ and hence $(u_{11} y_1 + u_{21} y_2) x = t$. This means that $x \,|_r\, t$ and therefore t is a gcrd of γ_1 and γ_2.

Moreover, $|\operatorname{lc} u_{21}| = |\operatorname{lc} g|$ since $|\operatorname{lc} u_{21}| = |\operatorname{lc} d_l| \cdots |\operatorname{lc} d_2| = |\operatorname{lc} d_2| \cdots |\operatorname{lc} d_l| = |\operatorname{lc} g|$. Now, it still remains to prove that $u_{21} \sim_J g$. From equation (3.6) we obtain $u_{21}\gamma_1 + u_{22}\gamma_2 = 0$, which is equivalent to $u_{21}\widetilde{g} + u_{22}g = 0$. Moreover, $\widetilde{g}$ and g are clearly right coprime and u_{21} and u_{22} are left coprime since they form a row of the unimodular matrix U. Thus, by Definition 2.3.1, it follows that $u_{21} \sim_J g$.

The case where $\deg \gamma_2 \geq \deg \gamma_1$ is analogous since, with the multiplication

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \gamma_2 \\ \gamma_1 \end{bmatrix},$$

we fall into the previous case, now with the expressions for γ_1 and γ_2 in (3.5) as well as the expressions for u_{21} and u_{22} in (3.7) being interchanged. It is clear that under these conditions the result still holds. □

Remark 3.1.9. Note that this result states that the well-known Diophantine equation

$$x\gamma_1 + y\gamma_2 = t$$

has a solution (x, y) that corresponds to a special unimodular completion

$$U = \begin{bmatrix} x & y \\ * & * \end{bmatrix}$$

lying in $Sl(2, \mathbb{H}[s])$. □

As in the commutative case, it is possible to obtain a triangular matrix from a given matrix in $\mathbb{H}^{n\times n}[s]$ by pre-multiplication by a matrix in $SL(n, \mathbb{H}[s])$. The result is formalized next. The proof is completely analogous to the one of [43, Theorem B.1.1] and is once more based on the Euclidian division algorithm.

Lemma 3.1.10. *For every $R \in \mathbb{H}^{n\times n}[s]$ there exists a matrix $U \in SL(n, \mathbb{H}[s])$ such that*

$$UR = T,$$

where $T \in \mathbb{H}^{n\times n}[s]$ is a triangular matrix.

Full row rank quaternionic (L-) polynomial matrices will play an important role in this thesis. The successive application of Lemma 3.1.8 allows to derive the following theorem, which states that it is always possible to reduce a quaternionic (L-) polynomial matrix R to an equivalent one of the form $\left[\begin{smallmatrix}\widetilde{R}\\0\end{smallmatrix}\right]$ with $\widetilde{R}$ frr.

Theorem 3.1.11. *Let $R \in \mathbb{H}^{g\times r}[s]$. Then there exists a unimodular matrix $U \in \mathbb{H}^{g\times g}[s]$ such that*

$$UR = \begin{bmatrix}\widetilde{R}\\0\end{bmatrix}$$

with $\widetilde{R} \in \mathbb{H}^{\tilde{g}\times r}[s]$ frr.

The next theorem concerns reduction to row proper form and is needed for the calculus of the solutions of quaternionic difference and differential equations performed in Section 4.3. Before stating it, we define the *degree of a row* (*row degree*) of a polynomial matrix as the maximum degree of the entries of that row.

Theorem 3.1.12. *Let $R(s) \in \mathbb{H}^{g\times r}[s]$ be frr. Then there exists $U(s) \in \mathbb{H}^{g\times g}[s]$ unimodular such that UR is row proper, i.e.,*

$$U(s)R(s) = D(s)\widetilde{R}_h + \widetilde{R}(s), \quad D(s) = \begin{bmatrix} s^{m_1} & & \\ & \ddots & \\ & & s^{m_g}\end{bmatrix},$$

where $\widetilde{R}_h \in \mathbb{H}^{g\times r}$ is frr and each row degree of $\widetilde{R}(s)$ is strictly smaller than the corresponding row degree of $D(s)\widetilde{R}_h$ (or equivalently of $D(s)$).

Proof. This proof is analogous to the one given in [13] for real case, with small adjustments. Let $R(s) = R_p s^p + \cdots + R_1 s + R_0 \in \mathbb{H}^{g\times r}[s]$ be frr. It is clear that the matrix $R(s)$ can be written in the form

$$R(s) = \begin{bmatrix} s^{m_1} & & \\ & \ddots & \\ & & s^{m_g}\end{bmatrix}\widetilde{R}_h + \widetilde{R}(s), \tag{3.8}$$

where $\widetilde{R}_h \in \mathbb{H}^{g\times r}$, $m_n \leq p$ and the degree of the n^{th} row of $\widetilde{R}(s)$ is smaller than m_n, $n = 1, \ldots, g$.

If $\widetilde{R}_h$ is frr the theorem follows with $U = I_g$.

Otherwise, there exists a non zero quaternionic vector $q = \begin{bmatrix} q_1 & q_2 & \cdots & q_g \end{bmatrix}$ such that $q\widetilde{R}_h = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}$.

Let the l^{th} row of $R(s)$ be of maximum degree among the ones to which correspond a non zero q_n. Then the polynomial vector

$$q(s) = \begin{bmatrix} q_1 s^{m_l - m_1} & \cdots & q_l & \cdots & q_g s^{m_l - m_g} \end{bmatrix}$$

satisfies

$$q(s) \begin{bmatrix} s^{m_1} & & \\ & \ddots & \\ & & s^{m_g} \end{bmatrix} \widetilde{R}_h = s^{m_l} \begin{bmatrix} q_1 & q_2 & \cdots & q_g \end{bmatrix} \widetilde{R}_h = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}.$$

Let $U_1(s) = \begin{bmatrix} \mathbf{e}_1 \,|\, \cdots \,|\, \mathbf{e}_{l-1} \,|\, q(s)^T \,|\, \mathbf{e}_{l+1} \,|\, \cdots \,|\, \mathbf{e}_g \end{bmatrix}^T$ and $R_1(s) = U_1(s)R(s)$, where $\mathbf{e}_l$ are the usual versors. Clearly $U_1(s)$ is a unimodular matrix.

It can easily be checked that the l^{th} row of $R_1(s)$ has degree smaller than m_l and all the other rows coincide with the corresponding rows of $R(s)$. Write $R_1(s)$ in the form (3.8). If $\widetilde{R}_{1h}$ is frr the proof is finished. Otherwise, repeat the whole procedure. In a finite number of steps, at most $\sum_{n=1}^{g} m_n$, the desired matrix is obtained. $\square$

3.2 The determinant of quaternionic polynomial matrices

As is well-known, determinants play an important role in the study of the stability of dynamical systems. However, although different types of determinants have been considered for matrices over $\mathbb{H}$, up to our knowledge, this work has not been extended for matrices over the polynomial ring $\mathbb{H}[s]$. This has motivated our search for a suitable notion of determinant for quaternionic polynomial matrices. The definition proposed here is inspired by Dieudonné's approach for matrices over the skew-field $\mathbb{H}$ (cf Section 1.2.2), but contains some necessary adaptations to the polynomial ring case. Indeed, the straightforward extension of the Dieudonné determinant to the polynomial case faces two major difficulties. First it is impossible to diagonalize a polynomial matrix $R \in \mathbb{H}^{n\times n}[s]$ as in Lemma 1.2.6, i.e., only multiplying on the left by a matrix $U \in SL(n, \mathbb{H}[s])$. Second it does not make sense to define the norm of a polynomial.

However, as stated in Lemma 3.1.10, given a matrix $R \in \mathbb{H}^{n\times n}[s]$ there exists a matrix $U \in SL(n, \mathbb{H}[s])$ such that $UR = T$, where T is a triangular polynomial matrix. Using an approach in some sense similar to the one of Dieudonné, we define a polynomial determinant for the quaternionic polynomial matrix R with basis on the diagonal elements of the triangular matrix T. First we extend the general definition of determinant given in Definition 1.2.5 to quaternionic polynomial matrices.

Definition 3.2.1. A function $d : \mathbb{H}^{n\times n}[s] \to \mathbb{H}[s]$ is said to be a *polynomial determinant* if it satisfies the following axioms:

(i) $d(A) = 0$ if and only if A has not full rank.

(ii) $d(AB) = d(A)d(B)$ for all $A, B \in \mathbb{H}^{n\times n}[s]$.

(iii) If $A' = B_{lm}(\alpha)A$, $\alpha \in \mathbb{H}[s]$, then $d(A') = d(A)$.

Definition 3.2.2. We define the function $\mathrm{Pdet}(\cdot) : \mathbb{H}^{n\times n}[s] \to \mathbb{R}[s]$ as follows. Let $R \in \mathbb{H}^{n\times n}[s]$. Let further $U \in SL(n, \mathbb{H}[s])$ be such that UR is upper triangular, i.e.,

$$UR = T = \begin{bmatrix} \gamma_1 & * & \cdots & * & * \\ 0 & \gamma_2 & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & * & * \\ \vdots & \vdots & \ddots & \gamma_{n-1} & * \\ 0 & 0 & \cdots & 0 & \gamma_n \end{bmatrix}. \tag{3.9}$$

Then

$$\mathrm{Pdet}(R) = \prod_{l=1}^{n} \gamma_l \overline{\gamma}_l.$$

Remark 3.2.3. If, in particular, R is a constant matrix, i.e., $R \in \mathbb{H}^{n\times n}$, $\mathrm{Pdet}(R) = \left[\mathrm{Ddet}(R)\right]^2$. Indeed, by Lemma 1.2.6, there exists a matrix $U \in SL(n, \mathbb{H})$ such that $UR = \mathrm{diag}(1, \dots, 1, \alpha)$, $\alpha \in \mathbb{H}$. Hence $\mathrm{Ddet}(R) = |\alpha|$ and $\mathrm{Pdet}(R) = \alpha\overline{\alpha} = |\alpha|^2$. Further, since by Proposition 1.2.11 $\left[\mathrm{Ddet}(R)\right]^2 = \mathrm{Sdet}(R)$, we conclude that in the constant case Pdet and Sdet coincide. □

Note that Definition 3.2.2 is well posed since, as we next show, if another triangular matrix $T' \neq T$, where T is the matrix defined in (3.9), is obtained by pre-multiplying the matrix R by $U' \in SL(n, \mathbb{H}[s])$, i.e., if $T' = U'R$, then the elements of the main diagonal of T', $\gamma'_1, \dots, \gamma'_n$, are such that

$$\prod_{l=1}^{n} \gamma'_l\overline{\gamma'}_l = \prod_{l=1}^{n} \gamma_l\overline{\gamma}_l. \tag{3.10}$$

Indeed, if R has not full rank, the same happens for every triangular matrix $\widetilde{T}$ such that $UR = \widetilde{T}$, for some $U \in SL(n, \mathbb{H}[s])$. This clearly implies that at least one of the diagonal elements of T is zero, and the same happens with T'. Therefore (3.10) holds since both sides of the equality are zero.

Let now R have full rank. Suppose first that $R(s)$ is triangular and 2×2, i.e.,

$$R = \begin{bmatrix} \gamma_1 & \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix}, \quad \text{with} \quad \gamma_1, \gamma_2 \neq 0.$$

Let $U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \in SL(2, \mathbb{H}[s])$ be such that

$$UR = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \gamma_1 & \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix} = \begin{bmatrix} u_{11}\gamma_1 & u_{11}\gamma_{12} + u_{12}\gamma_2 \\ u_{21}\gamma_1 & u_{21}\gamma_{12} + u_{22}\gamma_2 \end{bmatrix} = \begin{bmatrix} \gamma'_1 & \gamma'_{12} \\ 0 & \gamma'_2 \end{bmatrix} = T'.$$

Then $u_{21}\gamma_1 = 0$, i.e., $u_{21} = 0$ and therefore U is triangular. Taking into account that $U \in SL(2, \mathbb{H}[s])$, this implies that u_{11} and u_{22} are nonzero constants, i.e., $u_{11}, u_{22} \in \mathbb{H} \setminus \{0\}$. We next show that $|u_{11}||u_{22}| = 1$. Indeed, it is not difficult to see that there exists $V \in SL(2, \mathbb{H}[s])$ such that

$$VU = \begin{bmatrix} 1 & 0 \\ 0 & u_{11}u_{22} \end{bmatrix}.$$

Since $VU \in SL(2,\mathbb{H}[s])$ and is a constant matrix, this implies that $VU \in SL(2,\mathbb{H})$. Consequently the Dieudonné determinant of VU must be equal to 1, i.e.,

$$\mathrm{Ddet}(VU) = |u_{11}u_{22}| = |u_{11}||u_{22}| = 1.$$

Recall that, as mentioned in Proposition 2.1.3, for every $p, q \in \mathbb{H}[s]$, $p\overline{p} \in \mathbb{R}[s]$ and $\overline{p}\,\overline{q} = \overline{qp}$. Hence,

$$\gamma_1'\overline{\gamma'}_1\gamma_2'\overline{\gamma'}_2 = u_{11}\gamma_1\overline{\gamma}_1\overline{u_{11}}u_{22}\gamma_2\overline{\gamma}_2\overline{u_{22}} = \gamma_1\overline{\gamma}_1\gamma_2\overline{\gamma}_2|u_{11}|^2|u_{22}|^2 = \gamma_1\overline{\gamma}_1\gamma_2\overline{\gamma}_2.$$

If $R \in \mathbb{H}^{n\times n}[s]$ is triangular of arbitrary dimension and $U \in SL(n,\mathbb{H}[s])$ is such that $UR = T'$, with T' triangular, analogously to the previous case the matrix U must also be triangular and the product of the norms of its main diagonal elements is equal to 1. Since the proof of this fact is analogous to the 2×2 case, we omit it for the sake of simplicity. Thus, the equality (3.10) holds.

Finally, consider the case where R is not triangular. Let $U, U' \in SL(n,\mathbb{H}[s])$ be such that

$$UR = T, \quad U'R = T', \quad \text{with } \; T, T' \text{ triangular.}$$

Then

$$T' = U'R = U'U^{-1}UR = U''T,$$

where $U'' = U'U^{-1} \in SL(n,\mathbb{H}[s])$, and, by the previous case, we can once more conclude that the equality (3.10) holds.

Example 3.2.4. Let

$$R(s) = \begin{bmatrix} (s+2\mathbf{j})(s+\mathbf{j}) & (s+2\mathbf{j})(s+2\mathbf{k})+2s+3 \\ s+\mathbf{j} & (s+2\mathbf{k}) \end{bmatrix}.$$

Then

$$R = UT = \begin{bmatrix} s+2\mathbf{j} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} s+\mathbf{j} & s+2\mathbf{k} \\ 0 & 2s+3 \end{bmatrix},$$

where $U \in SL(2,\mathbb{H}[s])$. On the other hand,

$$R = U'T' = \begin{bmatrix} -\frac{1}{2}(s+2\mathbf{j}) & s+2+2\mathbf{j} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} -2(s+\mathbf{j}) & 3-4\mathbf{k} \\ 0 & s+\frac{3}{2} \end{bmatrix}.$$

Note that $U' \in SL(2, \mathbb{H}[s])$ since

$$U' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ s+2\mathbf{j} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

Therefore

$$\begin{aligned} \mathrm{Pdet}(R) &= (s+\mathbf{j})(\overline{s+\mathbf{j}})(2s+3)(\overline{2s+3}) \\ &= \big(-2(s+\mathbf{j})\big)\big(\overline{-2(s+\mathbf{j})}\big)\big(s+\tfrac{3}{2}\big)\big(\overline{s+\tfrac{3}{2}}\big) \\ &= (s^2+1)(2s+3)^2. \end{aligned}$$

□

In order to show that Pdet(.) is indeed a polynomial determinant in the sense of Definition 3.2.1, we first prove an auxiliary result that states that $\mathrm{Pdet}(R)$ is invariant with respect to the post-multiplication of R by a matrix $U \in SL(n, \mathbb{H}[s])$.

Lemma 3.2.5. *Let $M \in \mathbb{H}^{n\times n}[s]$ and $U \in SL(n, \mathbb{H}[s])$. Then*

$$\mathrm{Pdet}(MU) = \mathrm{Pdet}(M). \tag{3.11}$$

Proof. Since $U \in SL(n, \mathbb{H}[s])$ is a finite product of P_{lm} and $B_{lm}(\alpha)$ matrices, $\alpha \in \mathbb{H}[s]$, it is clearly enough to prove that

$$\mathrm{Pdet}(MS) = \mathrm{Pdet}(M)$$

if S is a P_{lm} or $B_{lm}(\alpha)$ matrix, $\alpha \in \mathbb{H}[s]$.

If M has not full rank, the equality (3.11) trivially holds.

If M has full rank, consider a matrix $V \in SL(n, \mathbb{H}[s])$ such that $T = VM$, where T is triangular, and consequently $\mathrm{Pdet}(M) = \mathrm{Pdet}(T)$. Then

$$\mathrm{Pdet}(MS) = \mathrm{Pdet}(VMS) = \mathrm{Pdet}(TS)$$

and it is therefore sufficient to show that $\mathrm{Pdet}(TS) = \mathrm{Pdet}(S)$.

Suppose first that $T \in \mathbb{H}^{2\times 2}[s]$ and is given by $T = \begin{bmatrix} \gamma_1 & \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix}$, with $\gamma_1, \gamma_2 \neq 0$, and hence

$$\mathrm{Pdet}(T) = \gamma_1\overline{\gamma}_1\gamma_2\overline{\gamma}_2. \tag{3.12}$$

In the sequel we show that $\mathrm{Pdet}(TS) = \mathrm{Pdet}(T)$, where $S = B_{12}(\alpha)$, $\alpha \in \mathbb{H}[s]$, or $S = P_{12}$. Note that it is not necessary to prove that $\mathrm{Pdet}(TS) = \mathrm{Pdet}(T)$ with $S = B_{21}(\alpha)$ since $B_{21}(\alpha) = P_{12}B_{12}(\alpha)P_{12}$. The first case is obvious because

$$TB_{12}(\alpha) = \begin{bmatrix} \gamma_1 & \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \gamma_1 & \gamma_1\alpha + \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix}.$$

On the other hand,

$$TP_{12} = \begin{bmatrix} \gamma_1 & \gamma_{12} \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \gamma_{12} & \gamma_1 \\ \gamma_2 & 0 \end{bmatrix} = T'.$$

By Lemma 3.1.8 there exists a matrix $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \in SL(n, \mathbb{H}[s])$ such that

$$VT' = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} \gamma_{12} & \gamma_1 \\ \gamma_2 & 0 \end{bmatrix} = \begin{bmatrix} \gamma_1' & v_{11}\gamma_1 \\ 0 & v_{21}\gamma_1 \end{bmatrix},$$

where γ_1' is a gcrd of γ_{12} and γ_2, and letting g be such that $\gamma_2 = g\gamma_1'$, $v_{21} \sim_J g$ and $|\operatorname{lc} v_{21}| = |\operatorname{lc} g|$. Therefore

$$\mathrm{Pdet}(TP_{12}) = \mathrm{Pdet}(T') = \gamma_1'\overline{\gamma'}_1 v_{21}\gamma_1\overline{v_{21}\gamma_1} = \gamma_1'\overline{\gamma'}_1 v_{21}\overline{v_{21}}\gamma_1\overline{\gamma}_1.$$

By Proposition 2.3.3, $v_{21}\overline{v_{21}} = g\overline{g}$, and thus, since $\gamma_2\overline{\gamma}_2 = g\gamma_1'\overline{g\gamma_1'} = \gamma_1'\overline{\gamma'}_1 g\overline{g}$, by (3.12)

$$\mathrm{Pdet}(TP_{12}) = \gamma_1'\overline{\gamma'}_1 v_{21}\overline{v_{21}}\gamma_1\overline{\gamma}_1 = \gamma_1'\overline{\gamma'}_1 g\overline{g}\gamma_1\overline{\gamma}_1 = \gamma_2\overline{\gamma}_2\gamma_1\overline{\gamma}_1 = \mathrm{Pdet}(T).$$

The case where $T \in \mathbb{H}^{n\times n}[s]$ has arbitrary dimension can be treated with basis on the previous one. In fact, if $S = B_{lm}(\alpha)$, $\alpha \in \mathbb{H}[s]$, the proof that $\mathrm{Pdet}(TS) = \mathrm{Pdet}(T)$ is analogous to the 2×2 case. Moreover, note that P_{lm} can be written as a product of matrices $P_{r(r+1)}$, i.e, matrices that are obtained from the identity by changing consecutive rows. Indeed,

$$P_{lm} = \left(\prod_{r=l}^{m-1} P_{r(r+1)}\right)\left(\prod_{r=2-m}^{-l} P_{(-r)(-r+1)}\right), \; l < m;$$

for instance

$$P_{14} = P_{12}P_{23}P_{34}P_{23}P_{12}.$$

The proof that $\mathrm{Pdet}(TP_{r(r+1)}) = \mathrm{Pdet}(T)$ is analogous to the 2×2 case and the result follows. $\square$

Having proved this auxiliary result, we are now able to show that Pdet satisfies the requirements of a polynomial determinant.

Proposition 3.2.6. Pdet *is a polynomial determinant, i.e.,*

(i) $\mathrm{Pdet}(R) = 0$ *if and only if* R *has not full rank.*

(ii) $\mathrm{Pdet}(RR') = \mathrm{Pdet}(R)\,\mathrm{Pdet}(R')$ *for all* $R, R' \in \mathbb{H}^{n\times n}[s]$.

(iii) If $R' = B_{lm}(\alpha)R$, $\alpha \in \mathbb{H}[s]$, *then* $\mathrm{Pdet}(R') = \mathrm{Pdet}(R)$.

Proof. *(i)* If R has not full rank and $UR = T$, for some $U \in SL(n, \mathbb{H}[s])$ and T a triangular matrix, then T has not full rank. Therefore one of it its main diagonal elements is zero and hence $\mathrm{Pdet}(T) = 0$. On the other hand, let $UR = T$, with U and T as in Definition 3.2.2. If $\mathrm{Pdet}(R) = 0$, then some entry γ_l on the main diagonal of T must be zero. This implies that the matrix T has not full rank, and hence R has not full rank.

(ii) Let $R, R' \in \mathbb{H}^{n\times n}[s]$ and $U, U' \in SL(n, \mathbb{H}[s])$ be such that

$$UR = T, \quad U'R' = T', \tag{3.13}$$

where T and T' are triangular matrices whose main diagonal elements are, respectively, γ_l and γ_l', $l = 1, \ldots, n$.

By definition we have

$$\mathrm{Pdet}(R) = \prod_{l=1}^{n} \gamma_l \overline{\gamma}_l \quad \text{and} \quad \mathrm{Pdet}(R') = \prod_{l=1}^{n} \gamma_l' \overline{\gamma'}_l.$$

Therefore

$$\mathrm{Pdet}(R)\,\mathrm{Pdet}(R') = \prod_{l=1}^{n} \gamma_l \overline{\gamma}_l \gamma_l' \overline{\gamma'}_l = \prod_{l=1}^{n} \gamma_l \gamma_l' \overline{\gamma'}_l \overline{\gamma}_l = \prod_{l=1}^{n} \gamma_l \gamma_l' \overline{\gamma_l \gamma_l'}. \tag{3.14}$$

Now, note that by (3.13)

$$URR' = TR' = TU'^{-1}T'. \tag{3.15}$$

Let $V \in SL(n, \mathbb{H}[s])$ be such that $TU'^{-1} = V\widetilde{T}$, with $\widetilde{T}$ triangular. It follows from Lemma 3.2.5 that

$$\mathrm{Pdet}(\widetilde{T}) = \mathrm{Pdet}(TU'^{-1}) = \mathrm{Pdet}(T); \tag{3.16}$$

moreover by (3.15)

$$V^{-1}URR' = \widetilde{T}T'.$$

Thus, since $V^{-1}U \in SL(n, \mathbb{H}[s])$

$$\mathrm{Pdet}(RR') = \mathrm{Pdet}(\widetilde{T}T') = \mathrm{Pdet}(\widetilde{T})\,\mathrm{Pdet}(T'),$$

taking into account that the main diagonal elements of $\widetilde{T}T'$ are the products of the corresponding main diagonal elements of $\widetilde{T}$ and T'. Finally, since from (3.16) $\mathrm{Pdet}(\widetilde{T}) = \mathrm{Pdet}(T)$, we conclude that

$$\mathrm{Pdet}(RR') = \mathrm{Pdet}(\widetilde{T})\,\mathrm{Pdet}(T') = \mathrm{Pdet}(T)\,\mathrm{Pdet}(T') = \mathrm{Pdet}(R)\,\mathrm{Pdet}(R').$$

(iii) By *(ii)*, $\mathrm{Pdet}(R') = \mathrm{Pdet}(B_{lm}(\alpha)R) = \mathrm{Pdet}\left(B_{lm}(\alpha)\right)\mathrm{Pdet}(R)$. The result follows since it is obvious that $\mathrm{Pdet}\left(B_{lm}(\alpha)\right) = 1$. □

In the sequel we investigate the connection between the zeros of $\mathrm{Pdet}\,R$ and the right eigenvalues of the companion matrix [28] associated to $R(s)$. This is relevant for the study of asymptotic stability of behavioral systems which is done in Section 5.3. We start by relating the right eigenvalues of a constant matrix $A \in \mathbb{H}^{n\times n}$ with the zeros of $\mathrm{Pdet}(sI - A)$.

Theorem 3.2.7. *Let $A \in \mathbb{H}^{n\times n}$. Then*

$$\lambda \in \sigma_r(A) \;\Leftrightarrow\; \lambda \textit{ is a zero of } \mathrm{Pdet}(sI - A).$$

Proof. Assume first that A is full rank (invertible). Then, as happens for real matrices, there exists an invertible matrix $S \in \mathbb{H}^{n\times n}$ such that $A' = S^{-1}AS$ is a companion matrix, i.e., has the form

$$A' = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \in \mathbb{H}^{n\times n}.$$

Note that $\sigma_r(A') = \sigma_r(A)$. Indeed, $A'v = v\lambda$ for some nonzero $v \in \mathbb{H}^n$ if and only if $Av' = v'\lambda$, where $v' = Sv$. On the other hand, $sI - A = sI - SA'S^{-1} = S(sI - A')S^{-1}$, which implies that $\operatorname{Pdet}(sI - A) = \operatorname{Pdet}(sI - A')$. Moreover, consider the matrices

$$P = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ p_1 & p_2 & \cdots & p_{l-1} & 1 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & s \\ s & 1 & \ddots & 0 & s^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s^{n-2} & s^{n-3} & \cdots & 1 & s^{n-1} \end{bmatrix},$$

where $p_l = s^{n-l} + a_{n-1}s^{n-(l+1)} + \cdots + a_{l+1}s + a_l \in \mathbb{H}[s]$, $l = 1, \ldots, n-1$. It is not difficult to check that P and $Q \in SL(n, \mathbb{H}[s])$ and that

$$P(sI - A')Q = \operatorname{diag}\big(-1, \ldots, -1, d(s)\big) = D,$$

with $d(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \in \mathbb{H}[s]$. Then,

$$\operatorname{Pdet}(sI - A') = \operatorname{Pdet}(P^{-1}DQ^{-1}) = \operatorname{Pdet}(P^{-1})\operatorname{Pdet}(D)\operatorname{Pdet}(Q^{-1}) = d\overline{d}.$$

Hence it suffices to prove that

$$\lambda \in \sigma_r(A') \;\Leftrightarrow\; \lambda \text{ is a zero of } d\overline{d}.$$

"$\Leftarrow$" Let λ be a zero of $d\overline{d}$. By Proposition 2.1.3, λ is a zero of $\overline{d}d$, i.e., $\big(\overline{d}d\big)(\lambda) = 0$. If $d(\lambda) = 0$, i.e., $\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$, it is immediate that $A'v = v\lambda$, with $v = \begin{bmatrix} 1 & \lambda & \cdots & \lambda^{n-1} \end{bmatrix}^T$. Hence $\lambda \in \sigma_r(A')$.

Suppose now that $d(\lambda) \neq 0$. Since, by hypothesis $\big(\overline{d}d\big)(\lambda) = 0$, by Proposition 2.2.5 there exists $\lambda' \in [\lambda]$ such that $\overline{d}(\lambda') = 0$. This implies, by Corollary 2.2.8, that there exists $\lambda'' \in [\lambda'] = [\lambda]$ such that $d(\lambda'') = 0$. From the previous case we conclude that $\lambda'' \in \sigma_r(A')$ which, by Proposition 1.2.17, implies that $\lambda \in \sigma_r(A')$.

" $\Rightarrow$ " Let $\lambda \in \sigma_r(A')$, i.e., $A'v = v\lambda$, for some nonzero $v = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}^T \in \mathbb{H}^n$. Then

$$\begin{cases} v_2 = v_1\lambda \\ \cdots \\ v_n = v_{n-1}\lambda \\ -a_0v_1 - a_1v_2 - \cdots - a_{n-1}v_n = v_n\lambda \end{cases} . \tag{3.17}$$

If $v_1 = 0$ it follows from (3.17) that $v_2 = \ldots = v_n = 0$ and thus we assume that $v_1 \neq 0$. Suppose first that $v_1 = 1$. In this case, from (3.17) we have that

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0,$$

i.e., $d(\lambda) = 0$. By Proposition 2.2.3 this implies that $\left(\overline{d}d\right)(\lambda) = 0$. If $v_1 \neq 1$, define $\widetilde{v} = vv_1^{-1} \in \mathbb{H}^n$ and $\widetilde{\lambda} = v_1\lambda v_1^{-1} \in \mathbb{H}$. Note that $\widetilde{v}_1 = 1$ and $\lambda \in [\widetilde{\lambda}]$. Then

$$A'v = v\lambda \;\Leftrightarrow\; A'vv_1^{-1} = vv_1^{-1}v_1\lambda v_1^{-1} \;\Leftrightarrow\; A'\widetilde{v} = \widetilde{v}\widetilde{\lambda}$$

and therefore $\widetilde{\lambda} \in \sigma_r(A')$. Analogously to the previous case, we conclude that $d(\widetilde{\lambda}) = 0$. By Lemma 2.2.10-4 this implies that $\left(\overline{d}d\right)(\nu) = 0$ for all $\nu \in [\widetilde{\lambda}]$ and hence $\left(\overline{d}d\right)(\lambda) = 0$.

Suppose now that A is not invertible. Let $V \in \mathbb{H}^{n\times n}$ be a change of coordinates that reduces A to its Jordan form $J = VAV^{-1}$, [55]. By the same arguments as before, it is clear that $\sigma_r(J) = \sigma_r(A)$ and $\operatorname{Pdet}(sI - J) = \operatorname{Pdet}(sI - A)$. Therefore, we will assume without loss of generality that $A = J$. Since A is not invertible it has a block diagonal form $A = \operatorname{diag}(N, \widetilde{A})$, where

$$N = \begin{bmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & * \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{H}^{r\times r} \quad \text{and} \quad \widetilde{A} \in \mathbb{H}^{(n-r)\times(n-r)} \;\text{ is invertible.}$$

Note that $\operatorname{Pdet}(sI - A) = s^{2r}\operatorname{Pdet}(sI - \widetilde{A})$. Hence, denoting by $\mathcal{N}\big(M(s)\big)$ the set of zeros of $\operatorname{Pdet}(M)$, $\mathcal{N}(sI - A) = \{0\} \cup \mathcal{N}(sI - \widetilde{A})$. On the other hand, $\sigma_r(A) =$

$\{0\} \cup \sigma_r(\widetilde{A})$. Since $\widetilde{A}$ is invertible, it follows from the first part that $\mathcal{N}(sI - \widetilde{A}) = \sigma_r(\widetilde{A})$, yielding the desired result. □

Remark 3.2.8. In [8] a similar result is given. Actually, [8, Theorem 8.5.1] states that if $d(s) \in \mathbb{H}[s]$ and A is the companion matrix of $d(s)$ then $\lambda \in \sigma_r(A)$ if and only if $d(\lambda) = 0$. However that result is not true. Indeed, if $d(s) = s^2 - (\mathbf{i} + \mathbf{j})s + \mathbf{k}$, the companion matrix of $d(s)$ is

$$A = \begin{bmatrix} 0 & 1 \\ -\mathbf{k} & \mathbf{i}+\mathbf{j} \end{bmatrix}.$$

It turns out that $\mathbf{i}$ is a right eigenvalue of A associated with the eigenvector $v = \left[\begin{smallmatrix} \mathbf{i}+\mathbf{j} \\ -1-\mathbf{k} \end{smallmatrix}\right]$, but $d(\mathbf{i}) = 2\mathbf{k} \neq 0$. Nevertheless, $\mathbf{i}$ is a zero of $\operatorname{Pdet}(sI - A) = (s^2+1)^2$. □

Corollary 3.2.9. *Let $R(s) = I_n s^m + R_{m-1}s^{m-1} + \cdots + R_1 s + R_0 \in \mathbb{H}^{n\times n}[s]$ and*

$$A = \begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_n \\ -R_0 & -R_1 & -R_2 & \cdots & -R_{m-1} \end{bmatrix} \in \mathbb{H}^{mn\times mn}$$

be the block companion matrix of R. Then

$$\lambda \in \sigma_r(A) \;\Leftrightarrow\; \big(\operatorname{Pdet} R\big)(\lambda) = 0.$$

Proof. Consider the matrix $(sI - A) \in \mathbb{H}^{mn\times mn}[s]$. Analogous to what happens in the proof of Theorem 3.2.7, there exist two matrices P and $Q \in SL(mn, \mathbb{H}[s])$ of the form

$$P = \begin{bmatrix} I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \\ P_1 & P_2 & \cdots & P_{l-1} & 1 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 0 & 0 & \cdots & 0 & I_n \\ I_n & 0 & \cdots & 0 & sI_n \\ sI_n & 1 & \ddots & 0 & s^2 I_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s^{m-2}I_n & s^{m-3}I_n & \cdots & 1 & s^{m-1}I_n \end{bmatrix},$$

where $P_l = s^{m-l}I_n + R_{m-1}s^{m-(l+1)} + \cdots + R_{l+1}s + R_l \in \mathbb{H}^{n\times n}[s]$, $l = 1, \ldots, m-1$, such that $P(sI - A)Q = D$, where D is the block diagonal matrix

$$D = \operatorname{diag}\big(-I, \ldots, -I, R(s)\big).$$

Then, it is clear that

$$\operatorname{Pdet}(sI - A) = \operatorname{Pdet}(P^{-1}DQ^{-1}) = \operatorname{Pdet}(P^{-1})\operatorname{Pdet}(D)\operatorname{Pdet}(Q^{-1}) = \operatorname{Pdet}(R),$$

and the result follows from Theorem 3.2.7. □

3.3 Polynomial complex adjoint matrix

The definition of the complex adjoint matrix of $R(s) \in \mathbb{H}^{g\times r}[s]$, $R^c(s)$, is analogous to the constant case. In fact, $R(s)$ may be uniquely written as $R(s) = R_1(s) + R_2(s)\mathbf{j}$, where $R_1(s)$, $R_2(s) \in \mathbb{C}^{g\times r}[s]$, and we define $R^c(s)$ as the $2g \times 2r$ complex polynomial matrix

$$R^c(s) = \begin{bmatrix} R_1(s) & R_2(s) \\ -\overline{R}_2(s) & \overline{R}_1(s) \end{bmatrix}. \tag{3.18}$$

Any complex matrix with the structure (3.18) is said to be a *complex adjoint matrix*.

The following theorem is a generalization of Theorem 1.2.3.

Theorem 3.3.1. *Let $A(s)$ and $B(s)$ be quaternionic polynomial matrices of suitable dimensions and, in case, invertible. Then*

1. $\big(A(s)B(s)\big)^c = A^c(s)B^c(s)$;
2. $\big(A(s) + B(s)\big)^c = A^c(s) + B^c(s)$;
3. $\big(A^{-1}\big)^c(s) = \big(A^c(s)\big)^{-1}$ *if A^{-1} exists.*

Proof. We will only prove the first statement. The other statements are proved by similar arguments.

Let $A(s) = A_1(s) + A_2(s)\mathbf{j} \in \mathbb{H}^{m\times n}[s]$, $B(s) = B_1(s) + B_2(s)\mathbf{j} \in \mathbb{H}^{n\times p}[s]$, where $A_1(s), A_2(s) \in \mathbb{C}^{m\times n}[s]$ and $B_1(s), B_2(s) \in \mathbb{C}^{n\times p}[s]$. Since, as referred in (1.2), for any

$z \in \mathbb{C}$, $\mathbf{j}z = \overline{z}\mathbf{j}$, we get that

$$\begin{aligned}
\big(A(s)B(s)\big)^c &= \Big[\big(A_1(s)+A_2(s)\mathbf{j}\big)\big(B_1(s)+B_2(s)\mathbf{j}\big)\Big]^c \\
&= \Big(A_1(s)B_1(s)+A_1(s)B_2(s)\mathbf{j}+A_2(s)\mathbf{j}B_1(s)+A_2(s)\mathbf{j}B_2(s)\mathbf{j}\Big)^c \\
&= \Big[\big(A_1(s)B_1(s)-A_2(s)\overline{B}_2(s)\big)+\big(A_1(s)B_2(s)+A_2(s)\overline{B}_1(s)\big)\mathbf{j}\Big]^c \\
&= \begin{bmatrix} A_1(s)B_1(s)-A_2(s)\overline{B}_2(s) & A_1(s)B_2(s)+A_2(s)\overline{B}_1(s) \\ -\Big(\overline{A_1(s)B_2(s)+A_2(s)\overline{B}_1(s)}\Big) & \overline{A_1(s)B_1(s)-A_2(s)\overline{B}_2(s)} \end{bmatrix} \\
&= \begin{bmatrix} A_1(s)B_1(s)-A_2(s)\overline{B}_2(s) & A_1(s)B_2(s)+A_2(s)\overline{B}_1(s) \\ -\overline{A}_1(s)\overline{B}_2(s)-\overline{A}_2(s)B_1(s) & \overline{A}_1(s)\overline{B}_1(s)-\overline{A}_2(s)B_2(s) \end{bmatrix} \\
&= \begin{bmatrix} A_1(s) & A_2(s) \\ -\overline{A}_2(s) & \overline{A}_1(s) \end{bmatrix} \begin{bmatrix} B_1(s) & B_2(s) \\ -\overline{B}_2(s) & \overline{B}_1(s) \end{bmatrix} \\
&= A^c(s)B^c(s)
\end{aligned}$$

□

Quaternionic polynomial matrices share many algebraic properties with their complex adjoint matrices, as we show in the following statements.

Lemma 3.3.2. *A quaternionic polynomial matrix R has frr if and only if R^c has frr. More generally, for every quaternionic polynomial matrix R,* $\operatorname{rank} R = n$ *if and only if* $\operatorname{rank} R^c = 2n$.

Proof. Let $R = R_1 + R_2\mathbf{j} \in \mathbb{H}^{g\times r}[s]$. First we prove that R has frr if and only if R^c has frr.

"Only if" part. Suppose that R has not frr. Then there exists a nonzero row vector $x \in \mathbb{H}^{1\times g}[s]$ such that $xR = 0$, hence, by Theorem 3.3.1-1, $x^cR^c = 0$ with $x^c \neq 0$, i.e., R^c has not frr.

"If" part. Suppose that R^c has not frr. Then there exists a nonzero complex polynomial row vector $y = \begin{bmatrix} y_1 & y_2 \end{bmatrix}$, with y_1, $y_2 \in \mathbb{C}^{1\times g}[s]$ such that

$$yR^c = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ -\overline{R}_2 & \overline{R}_1 \end{bmatrix} = 0. \tag{3.19}$$

Define $x \in \mathbb{H}^{1\times g}[s]$ as $x = y_1 + y_2\mathbf{j}$. It follows from equation (3.19) that

$$x^c R^c = \begin{bmatrix} y_1 & y_2 \\ -\overline{y}_2 & \overline{y}_1 \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ -\overline{R}_2 & \overline{R}_1 \end{bmatrix} = 0,$$

which is equivalent to $xR = 0$. Since $x \neq 0$, R has not frr.

Consider now the general case. If $R \in \mathbb{H}^{g\times r}[s]$ has frr the result follows from the first part. If R has not frr, by Theorem 3.1.11, there exists a unimodular matrix $U \in \mathbb{H}^{g\times g}[s]$ such that

$$UR = \begin{bmatrix} \widetilde{R} \\ 0 \end{bmatrix} \tag{3.20}$$

with $\widetilde{R} \in \mathbb{H}^{n\times r}[s]$ frr; clearly $\operatorname{rank} R = \operatorname{rank} \widetilde{R} = n$. By the first part we have that $\widetilde{R}^{\,c} \in \mathbb{H}^{2n\times 2r}[s]$ has frr and therefore $\operatorname{rank} \widetilde{R}^{\,c} = 2n$. Since, by (3.20) and Theorem 3.3.1-1,

$$\begin{bmatrix} \widetilde{R} \\ 0 \end{bmatrix}^c = (UR)^c = U^c R^c,$$

$\operatorname{rank} \left[\begin{smallmatrix} \widetilde{R} \\ 0 \end{smallmatrix}\right]^c = \operatorname{rank} R^c$, thus, since $\operatorname{rank} \left[\begin{smallmatrix} \widetilde{R} \\ 0 \end{smallmatrix}\right]^c = \operatorname{rank} \widetilde{R}^{\,c}$ we can conclude that $\operatorname{rank} R^c = \operatorname{rank} \widetilde{R}^{\,c} = 2n$. □

The next result is essential for the characterization of equivalent kernel representations of quaternionic behavioral systems that will be presented in Section 4.2.1.

Proposition 3.3.3. *Given two quaternionic polynomial matrices A and B, if the equation*

$$A^c = MB^c \tag{3.21}$$

holds for some complex polynomial matrix M, then there exists a quaternionic polynomial matrix T such that $A = TB$. Moreover, if B is frr then $M = T^c$.

Proof. Let $A = A_1 + A_2\mathbf{j}$, $B = B_1 + B_2\mathbf{j}$, and $M = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}$. Then,

$$A^c = MB^c \quad \Leftrightarrow \quad \begin{bmatrix} A_1 & A_2 \\ -\overline{A}_2 & \overline{A}_1 \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ -\overline{B}_2 & \overline{B}_1 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} A_1 & = & T_1B_1 - T_2\overline{B}_2 \\ A_2 & = & T_1B_2 + T_2\overline{B}_1 \\ -\overline{A}_2 & = & T_3B_1 - T_4\overline{B}_2 \\ \overline{A}_1 & = & T_3B_2 + T_4\overline{B}_1 \end{cases} \tag{3.22}$$

Let $T = T_1 + T_2\mathbf{j}$. By (3.22) it follows that

$$T^cB^c = \begin{bmatrix} T_1 & T_2 \\ -\overline{T}_2 & \overline{T}_1 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ -\overline{B}_2 & \overline{B}_1 \end{bmatrix} = \begin{bmatrix} T_1B_1 - T_2\overline{B}_2 & T_1B_2 + T_2\overline{B}_1 \\ -\overline{T}_2B_1 - \overline{T}_1\overline{B}_2 & -\overline{T}_2B_2 + \overline{T}_1\overline{B}_1 \end{bmatrix} = A^c$$

and so $A = TB$. Now suppose that B is frr. By Lemma 3.3.2 we have that B^c is also a frr matrix and, since $(M - T^c)B^c = 0$, we obtain that $M = T^c$. □

Corollary 3.3.4. *Let $U \in \mathbb{H}^{g\times g}[s]$. Then U is unimodular if and only if $U^c \in \mathbb{C}^{2g\times 2g}[s]$ is unimodular.*

Proof. "Only if" part. Let $U \in \mathbb{H}^{g\times g}[s]$ be unimodular. Then there exists $V \in \mathbb{H}^{g\times g}[s]$ such that

$$\begin{aligned} VU = I_g \quad &\Leftrightarrow \quad (VU)^c = I_g^c \\ &\Leftrightarrow \quad V^cU^c = I_{2g} \end{aligned}$$

i.e., U^c is unimodular.

"If" part. If U^c is unimodular, there exists $W \in \mathbb{C}^{2g\times 2g}[s]$ such that $I_{2g} = WU^c$. From Proposition 3.3.3 we conclude that there exists V such that $V^c = W$. Therefore $I_{2g} = V^cU^c$ and hence $VU = I_g$, i.e, U is unimodular. □

In section 1.2.2, according to [47], the Study determinant, $\mathrm{Sdet}(A)$, of a quaternionic matrix $A \in \mathbb{H}^{n\times n}$ was defined in terms of its complex adjoint matrix A^c. We extend this notion to the polynomial case, i.e., we define the Study determinant of any $R(s) \in \mathbb{H}^{n\times n}[s]$ as $\mathrm{Sdet}\big(R(s)\big) = \det\big(R^c(s)\big)$.

It was shown, in Remark 3.2.3, that Pdet and Sdet coincide in the constant case. The following theorem states that they also coincide in the polynomial case.

Theorem 3.3.5. *For every* $R \in \mathbb{H}^{n\times n}[s]$, $\mathrm{Pdet}(R) = \mathrm{Sdet}(R)$.

Proof. If R has not full rank the result is obvious, since both $\mathrm{Pdet}(R)$ and $\mathrm{Sdet}(R)$ are zero. Suppose then that R has full rank. Let $U \in SL(n, \mathbb{H}[s])$ be such that $UR = T$, where T is a triangular matrix with main diagonal elements $\gamma_1, \ldots, \gamma_n$. Then, by definition,

$$\mathrm{Pdet}(R) = \prod_{l=1}^{n} \gamma_l \overline{\gamma}_l.$$

On the other hand, we have that $U^c R^c = T^c$ and therefore $\det(R^c) = \det(T^c)$. It is not difficult to check that the matrix T^c is equivalent to a block triangular matrix with main diagonal blocks $\gamma_1^c, \ldots, \gamma_n^c$ and then

$$\det(R^c) = \prod_{l=1}^{n} \det\left(\gamma_l^c\right).$$

The proof is completed once we show that $\det\left(\gamma^c\right) = \gamma\overline{\gamma}$, for every $\gamma \in \mathbb{H}[s]$. Indeed, if $\gamma = \alpha + \beta\mathbf{j}$, with $\alpha, \beta \in \mathbb{C}[s]$, then

$$\gamma^c = \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix} \in \mathbb{C}^{2\times 2}[s].$$

By direct calculation, $\gamma\overline{\gamma} = \alpha\overline{\alpha} + \beta\overline{\beta} = \det(\gamma^c)$. $\square$

This result allows us to give an alternative definition of the characteristic polynomial of a quaternionic matrix in terms of the determinant Pdet. Indeed, in Proposition 1.2.10 we have defined the characteristic polynomial of a matrix $A \in \mathbb{H}^{n\times n}$ as $F_A(s) = \det(sI_{2n} - A^c) \in \mathbb{C}[s]$. Moreover, by Theorem 3.3.1, $sI_{2n} - A^c = (sI_n - A)^c$ and hence, by definition, $\mathrm{Sdet}(sI_n - A) = \det\big((sI_n - A)^c\big) = F_A(s)$. Thus, by Theorem 3.3.5, we conclude that the characteristic polynomial of A is also given by

$$F_A(s) = \mathrm{Pdet}(sI_n - A) \in \mathbb{R}[s].$$

This yields a more natural expression which could be taken as a new definition of characteristic polynomial.

3.4 Quaternionic Smith form

The real/complex Smith form plays an important role in the study of behavioral systems over commutative fields, in particular in the characterization of controllability and stability [25, 43]. As we will see in Chapter 5, the same happens for quaternionic systems. For this reason we dedicate this section to a detailed analysis of the quaternionic Smith form.

The main original result of this section is the characterization of the complex Smith form of complex adjoint matrices and its relation to the quaternionic Smith form of the corresponding quaternionic matrix. The Smith form for real and complex (L-) polynomial matrices has been already defined and deeply characterized by several authors (see [15, 28]). Before recalling its usual definition we introduce the following notation.

Notation 3.4.1. By $\mathrm{diag}(a_1, \ldots, a_n)$ we mean a (not necessarily square) matrix with suitable dimensions whose first n elements on the main diagonal are $a_1, \ldots, a_n$ and all the other entries are zero, i.e., $\mathrm{diag}(a_1, \ldots, a_n) \in \mathcal{M}^{g \times r}$ stands for the matrix (over $\mathcal{M}$)

$$\begin{array}{rc}
 & \left[\begin{array}{ccc|c} a_1 & & & \\ & \ddots & & 0 \\ & & a_n & \\ \hline & 0 & & 0 \end{array}\right] \\
g-n \text{ rows } \Big\{ & \hphantom{a_1\ \ddots\ a_n}\underbrace{\hphantom{0}}_{r-n \text{ columns}}
\end{array}$$

Theorem 3.4.2. [15, 25, 28, 43] *Let $R(s,s^{-1}) \in \mathbb{C}^{g\times r}[s,s^{-1}]$. Then there exist L-polynomial unimodular matrices $U(s,s^{-1})$ and $V(s,s^{-1})$ such that*

$$U(s,s^{-1})R(s,s^{-1})V(s,s^{-1}) = \Gamma(s) = \mathrm{diag}\left(\gamma_1(s),\dots,\gamma_n(s)\right) \in \mathbb{C}^{g\times r}[s],$$

where n is the rank of R, γ_l, $l=1,\dots,n$, are monic polynomials with nonzero independent term and $\gamma_l \mid \gamma_{l+1},\ l=1,\dots,n-1$.

The matrix $\Gamma(s)$ is said to be a *complex Smith form* of R.

If $R \in \mathbb{C}^{g\times r}[s]$ we can not guarantee that the polynomials $\gamma_l(s)$ have nonzero independent term.

The definition of Smith form for quaternionic (L-) polynomial matrices can also be found in literature (see [8, 17, 23]). While in the commutative case each entry on the main diagonal of the Smith form must be a divisor of the previous one, in the quaternionic case that entry is required to be a total divisor of the previous one. The quaternionic version of Theorem 3.4.2 is given next.

Theorem 3.4.3. [23] *Let $R(s,s^{-1}) \in \mathbb{H}^{g\times r}[s,s^{-1}]$ with rank n. Then there exist quaternionic L-polynomial unimodular matrices $U(s,s^{-1})$ and $V(s,s^{-1})$ such that*

$$U(s,s^{-1})R(s,s^{-1})V(s,s^{-1}) = \Gamma(s) = \mathrm{diag}\left(\gamma_1(s),\dots,\gamma_n(s)\right) \in \mathbb{H}^{g\times r}[s],$$

where γ_l, $l=1,\dots,n$, are monic polynomials with nonzero independent term and $\gamma_l \parallel \gamma_{l+1},\ l=1,\dots,n-1$.

The matrix $\Gamma(s)$ is said to be a *quaternionic Smith form* of R.

Once more, if $R \in \mathbb{H}^{g\times r}[s]$ we can not guarantee that the polynomials $\gamma_l(s)$ have nonzero independent term.

Proof. This proof is similar to the ones given in [23, Theorem 3.16] and in [43, Theorem 2.5.15]. We will only prove the case where R has fcr. The general case is proved analogously.

The proof is constructive. Suppose first that R is square and assume that R is nonzero. Multiplying R by a suitable unimodular matrix $U_1(s,s^{-1}) \in \mathbb{H}^{g\times g}[s,s^{-1}]$ we have that $R \in \mathbb{H}^{g\times g}[s]$. Apply row and column permutations so as to achieve that the nonzero element of minimal degree of R appears in the $(1,1)$ position, i.e., to obtain a matrix of the following form:

$$\begin{bmatrix} r_{11}(s) & r_{12}(s) & \cdots & r_{1g}(s) \\ r_{21}(s) & r_{22}(s) & \cdots & r_{2g}(s) \\ \vdots & \vdots & \ddots & \vdots \\ r_{g1}(s) & r_{g2}(s) & \cdots & r_{gg}(s) \end{bmatrix}, \tag{3.23}$$

where $r_{ln}(s) \in \mathbb{H}[s]$ and $\deg r_{11}(s) \leq \deg r_{ln}(s)$, $l,n = 1,\ldots,g$. Use the element $r_{11}(s)$ to carry out right division with remainder on the first column, i.e.,

$$r_{l1}(s) = d_{l1}(s)r_{11}(s) + x_{l1}(s), \quad \deg x_{l1}(s) < \deg r_{11}(s) \text{ or } x_{l1}(s) = 0,\ l = 2,\ldots,g.$$

Pre-multiplication the matrix given in equation (3.23) by the unimodular matrix

$$U(s) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -d_{21}(s) & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -d_{g1}(s) & 0 & \cdots & 1 \end{bmatrix}$$

we obtain the matrix

$$\begin{bmatrix} r_{11}(s) & r_{12}(s) & \cdots & r_{1g}(s) \\ x_{21}(s) & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ x_{g1}(s) & * & \cdots & * \end{bmatrix}, \tag{3.24}$$

Analogously, use the element $r_{11}(s)$ to carry out left division with remainder on the first row (which corresponds to post-multiplication by a unimodular matrix). Repeat the whole procedure as long as the first row or column contains at least two nonzero elements. Since degrees are bounded from bellow, this implies that within a finite number of steps we obtain a matrix of the following form:

$$\begin{bmatrix} * & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}. \tag{3.25}$$

Now, either the $(1,1)$ element in (3.25) is a total divisor of all the other elements in the matrix, or there exists a column that contains an element that is not a total multiple of the $(1,1)$ element. If the latter is true, add this column to the first column of (3.25) and repeat the previous procedure all over again.

Again after a finite number of steps we obtain a matrix of the form (3.25), but with a $(1,1)$ element of strictly smaller degree. As long as there is an element in the matrix that is not a total multiple of the $(1,1)$ element, we repeat this process.

As a consequence, in a finite number of steps, we obtain a matrix of the form (3.25) where the $(1,1)$ element is a total divisor of all the other elements. Moving on to the $(g-1)\times(g-1)$ right-lower submatrix and applying the whole procedure to that matrix, we obtain a matrix of the form

$$\begin{bmatrix} * & 0 & \cdots & \cdots & 0 \\ 0 & * & 0 & \cdots & 0 \\ \vdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix}, \tag{3.26}$$

where the $(2,2)$ element is a total divisor of all the elements of the $(g-2)\times(g-2)$ right-lower matrix. Note that the $(1,1)$ element is still a total divisor of all the other elements of the matrix (3.26). Continuing in this way, we obtain a diagonal matrix with the desired total division properties. If all its elements are monic polynomials with nonzero independent term, then that matrix is a quaternionic Smith form of R.

Otherwise, multiplying that diagonal matrix by a suitable unimodular matrix

$$U_2(s,s^{-1}) = \mathrm{diag}(\alpha_1 s^{m_1}, \ldots, \alpha_g s^{m_g}) \in \mathbb{H}^{g\times g}[s,s^{-1}], \quad \alpha_l \in \mathbb{H},\ l = 1,\ldots,g$$

we obtain a matrix where all its elements are monic polynomials with nonzero independent term and that, clearly, still has the total division properties.

If $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ with $g > r$, by Theorem 3.1.11, there exists a unimodular matrix U such that $UR = \left[\begin{smallmatrix}\widetilde{R}\\ 0\end{smallmatrix}\right]$, with $\widetilde{R}$ square. Thus, as proved above, there exist unimodular matrices $\widetilde{U}$ and $\widetilde{V}$ such that $\widetilde{U}\widetilde{R}\widetilde{V} = \widetilde{\Gamma}(s)$, where $\widetilde{\Gamma}$ is a quaternionic Smith form of $\widetilde{R}$. Therefore

$$\begin{bmatrix}\widetilde{U} & 0\\ 0 & I\end{bmatrix} UR\widetilde{V} = \begin{bmatrix}\widetilde{\Gamma}\\ 0\end{bmatrix},$$

yielding $\begin{bmatrix}\widetilde{\Gamma}\\ 0\end{bmatrix} = \Gamma$ as a quaternionic Smith form of R. $\square$

Note that, according to the previous theorem, the quaternionic Smith forms of a quaternionic L-polynomial matrix $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ are required to be quaternionic polynomial (and not L-polynomial) matrices, i.e., they must be elements of $\mathbb{H}^{g\times r}[s]$.

The following example shows that a quaternionic Smith form of a complex matrix does not coincide with its complex Smith form.

Example 3.4.4. Consider

$$R = \begin{bmatrix} s+\mathbf{i} & 0\\ 0 & s+\mathbf{i}\end{bmatrix}.$$

In the complex case this polynomial matrix is a complex Smith form (see Theorem 3.4.2). However, as we have shown in Remark 2.2.18, $(s+\mathbf{i}) \not\|\ (s+\mathbf{i})$ and therefore, R is not a quaternionic Smith form. A simple calculation shows that

$$\Gamma = URV = \begin{bmatrix} 1 & 0\\ 0 & s^2+1\end{bmatrix},$$

where

$$U = \begin{bmatrix} \frac{\mathbf{k}}{2} & \frac{\mathbf{i}}{2}\\ \mathbf{j}s+\mathbf{k} & s-\mathbf{i}\end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -\mathbf{j} & -\frac{\mathbf{k}}{2}(s+\mathbf{i})\\ -1 & \frac{\mathbf{i}}{2}(s-\mathbf{i})\end{bmatrix}$$

are unimodular polynomial matrices, is a quaternionic Smith form of R. $\square$

One main difference between the complex and the quaternionic case is that while the complex Smith form is unique, the quaternionic Smith form, in general, is not unique. This is illustrated in the next example.

Example 3.4.5. Consider the quaternionic polynomial matrix

$$R(s) = \begin{bmatrix} 1 & 0 \\ 0 & s+\mathbf{j} \end{bmatrix}.$$

It is clear that the matrix R is a quaternionic Smith form of itself. However,

$$\Gamma(s) = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & s+\mathbf{j} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\mathbf{i} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & s-\mathbf{j} \end{bmatrix}$$

is also a quaternionic Smith form of R and $R(s) \neq \Gamma(s)$. □

In this context, it is natural to ask which is the source of non uniqueness of the quaternionic Smith form. This problem has first been studied by Nakayama [34]. Before presenting Nakayama's result, we first state a simpler one.

Let $R \in \mathbb{H}^{g\times r}[s]$ and $\Gamma(s) = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ be a quaternionic Smith form of R. A sufficient condition for a matrix $\Gamma'(s) = \operatorname{diag}(\gamma'_1, \dots, \gamma'_n)$ to be also a quaternionic Smith form of R is given next.

Lemma 3.4.6. *Let $R \in \mathbb{H}^{g\times r}[s]$ and $\Gamma(s) = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ be a quaternionic Smith form of R. If*

$$\gamma'_l \sim \gamma_l, \quad \forall\, l = 1, \dots, n, \tag{3.27}$$

then $\Gamma'(s) = \operatorname{diag}(\gamma'_1, \dots, \gamma'_n)$ is a quaternionic Smith form of R.

Proof. Assume that (3.27) holds. Then, for each $l = 1, \dots, n$, there exists $\alpha_l \in \mathbb{H} \setminus \{0\}$ such that $\gamma'_l(s) = \alpha_l \gamma_l(s) \alpha_l^{-1}$. Hence,

$$\Gamma'(s) = \operatorname{diag}(\alpha_1, \dots, \alpha_n)\Gamma(s)\operatorname{diag}(\alpha_1^{-1}, \dots, \alpha_n^{-1}).$$

This means that $\Gamma'(s)$ is equivalent to $\Gamma(s)$ and, in turn, equivalent to $R(s)$.

It is obvious that the polynomials γ_l', $l = 1, \dots, n$, are monic. Moreover, by hypothesis, $\gamma_{l+1} \parallel \gamma_l$, and therefore, by (3.27) and the definition of total divisor we have that $\gamma_{l+1}' \parallel \gamma_l'$.

Thus, $\Gamma'(s)$ is a quaternionic Smith form of R. □

Unfortunately, the previous condition is not necessary, i.e., if $\Gamma(s) = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ and $\Gamma'(s) = \operatorname{diag}(\gamma_1', \dots, \gamma_n')$ are two quaternionic Smith forms of a quaternionic polynomial matrix R, this does not imply that $\gamma_l' \sim \gamma_l$, $\forall\, l = 1, \dots, n$.

Example 3.4.7. Let

$$R(s) = \begin{bmatrix} s - \mathbf{i} & 0 \\ 0 & s - 2\mathbf{j} \end{bmatrix} \in \mathbb{H}^{2\times 2}[s].$$

We will show that the matrices

$$\Gamma(s) = \begin{bmatrix} 1 & 0 \\ 0 & \left(s + \frac{3\mathbf{i}+4\mathbf{j}}{5}\right)(s - 2\mathbf{j}) \end{bmatrix} \quad \text{and} \quad \Gamma'(s) = \begin{bmatrix} 1 & 0 \\ 0 & (s - \mathbf{i})(s - 2\mathbf{j}) \end{bmatrix}$$

are both quaternionic Smith forms of R. However, as seen in Example 2.3.2,

$$\left(s + \tfrac{3\mathbf{i}+4\mathbf{j}}{5}\right)(s - 2\mathbf{j}) \not\sim (s - \mathbf{i})(s - 2\mathbf{j}).$$

Consider the quaternionic polynomial matrices

$$U(s) = \frac{1}{5}\begin{bmatrix} \mathbf{i} - 2\mathbf{j} & 2\mathbf{j} - \mathbf{i} \\ -5s - 8\mathbf{i} + 6\mathbf{j} & 5s + 3\mathbf{i} + 4\mathbf{j} \end{bmatrix} \quad \text{and} \quad V(s) = \frac{1}{5}\begin{bmatrix} 5 & (\mathbf{i} - 2\mathbf{j})s - 4 - 2\mathbf{k} \\ 5 & (\mathbf{i} - 2\mathbf{j})s + 1 - 2\mathbf{k} \end{bmatrix}.$$

These matrices are unimodular since their inverses

$$U^{-1}(s) = \begin{bmatrix} s - \mathbf{i} & \frac{1}{5}\mathbf{i} - \frac{2}{5}\mathbf{j} \\ s - 2\mathbf{j} & \frac{1}{5}\mathbf{i} - \frac{2}{5}\mathbf{j} \end{bmatrix}$$

and

$$V^{-1}(s) = \frac{1}{5}\begin{bmatrix} (\mathbf{i} - 2\mathbf{j})s + 1 - 2\mathbf{k} & (-\mathbf{i} + 2\mathbf{j})s + 4 - 2\mathbf{k} \\ -5 & 5 \end{bmatrix}$$

are also polynomial. Moreover, $\Gamma = URV$, which implies that the matrix Γ is a quaternionic Smith form of R.

On the other hand, the matrix Γ' is also a quaternionic Smith form of R since $\Gamma' = U'RV'$, where

$$U'(s) = \begin{bmatrix} \frac{5}{9} & \frac{\mathbf{i}+2\mathbf{j}}{3} \\ -\frac{(\mathbf{i}+2\mathbf{j})s+4-2\mathbf{k}}{3} & -\mathbf{i}+s \end{bmatrix} \quad \text{and} \quad V'(s) = \begin{bmatrix} -\frac{3\mathbf{i}+6\mathbf{j}}{5} & -s+2\mathbf{j} \\ 1 & \frac{(-\mathbf{i}+2\mathbf{j})s-1+2\mathbf{k}}{3} \end{bmatrix}$$

are unimodular matrices with inverses

$$U'^{-1}(s) = \begin{bmatrix} \frac{-(3\mathbf{i}+6\mathbf{j})s-3+6\mathbf{k}}{5} & -1 \\ s-2\mathbf{j} & -\frac{\mathbf{i}+2\mathbf{j}}{3} \end{bmatrix} \quad \text{and} \quad V'^{-1}(s) = \begin{bmatrix} \frac{5}{9}(s-\mathbf{i}) & \frac{(\mathbf{i}+2\mathbf{j})s+4-2\mathbf{k}}{3} \\ -\frac{\mathbf{i}+2\mathbf{j}}{3} & 1 \end{bmatrix}.$$

$\square$

In [34] the following necessary condition in terms of the J-similarity is given.

Theorem 3.4.8. [23, 34] *If*

$$\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \quad \textit{and} \quad \Gamma' = \operatorname{diag}(\gamma'_1, \ldots, \gamma'_n)$$

are two quaternionic Smith forms of a quaternionic polynomial matrix $R(s)$ then

$$\gamma_l \sim_J \gamma'_l, \quad l = 1, \ldots, n.$$

Combining this Lemma with Proposition 2.3.3 it is possible to give the following alternative necessary condition.

Corollary 3.4.9. *If*

$$\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \quad \textit{and} \quad \Gamma' = \operatorname{diag}(\gamma'_1, \ldots, \gamma'_n)$$

are two quaternionic Smith forms of a quaternionic matrix R, then

$$\gamma_l \overline{\gamma}_l = \gamma'_l \overline{\gamma'}_l, \quad l = 1, \ldots, n.$$

In [32] it is shown that the condition stated in Theorem 3.4.8 is not sufficient when $\operatorname{rank} R(s) = 1$. However, if $\operatorname{rank} R(s) \geq 2$, one can prove that the condition is indeed sufficient, [17]. This result is formalized next.

Theorem 3.4.10. [17] *Let* $R \in \mathbb{H}^{g\times r}[s]$ *with* $\operatorname{rank} R \geq 2$ *and* $\Gamma(s) = \operatorname{diag}(\gamma_1, \ldots, \gamma_n)$ *be a quaternionic Smith form of* R. *If*

$$\gamma_l' \sim_J \gamma_l, \quad \forall\, l = 1, \ldots, n,$$

then $\Gamma'(s) = \operatorname{diag}(\gamma_1', \ldots, \gamma_n')$ *is a quaternionic Smith form of* R.

Remark 3.4.11. Note that the reciprocal of Corollary 3.4.9 is not true even when $\operatorname{rank} R \geq 2$. For instance, let

$$R = \Gamma = \begin{bmatrix} 1 & 0 \\ 0 & s^2+1 \end{bmatrix} \in \mathbb{H}^{2\times 2}[s] \quad \text{and} \quad \Gamma' = \begin{bmatrix} 1 & 0 \\ 0 & (s+\mathbf{i})^2 \end{bmatrix} \in \mathbb{H}^{2\times 2}[s].$$

It can easily be checked that

$$(s^2+1)\overline{(s^2+1)} = (s^2+1)^2 = (s+\mathbf{i})^2\overline{(s+\mathbf{i})^2}.$$

Obviously, Γ is a quaternionic Smith form of R. On the other hand, the complex Smith forms of Γ^c and $(\Gamma')^c$ are, respectively,

$$\operatorname{diag}\left(1, 1, s^2+1, s^2+1\right) \quad \text{and} \quad \operatorname{diag}\left(1, 1, 1, \left(s^2+1\right)^2\right).$$

Since they are different we have that Γ^c and $(\Gamma')^c$ are not equivalent which implies that neither are Γ and Γ'. Hence Γ' cannot be a quaternionic Smith form of R. $\square$

In the sequel, we will study the relation between the quaternionic Smith form of a matrix $R \in \mathbb{H}^{g\times r}[s]$ and the complex Smith form of $R^c \in \mathbb{C}^{2g\times 2r}[s]$. We will as well characterize the special structure of the complex Smith form of polynomial complex adjoint matrices. First we give an auxiliary result.

Proposition 3.4.12. *Let* $q \in \mathbb{H}[s]$ *be monic. Then the complex Smith form of* q^c *is*

$$\operatorname{diag}\left(\mathcal{F}_q, \mathcal{M}_q\right).$$

Proof. Let $q = \mathcal{F}_q\mathcal{Q}_q$ and suppose that $\mathcal{Q}_q = q_1 + q_2\mathbf{j}$ for some $q_1, q_2 \in \mathbb{C}[s]$. Then $q = \mathcal{F}_q q_1 + \mathcal{F}_q q_2\mathbf{j}$ and therefore

$$q^c = \begin{bmatrix} \mathcal{F}_q q_1 & \mathcal{F}_q q_2 \\ -\mathcal{F}_q\overline{q}_2 & \mathcal{F}_q\overline{q}_1 \end{bmatrix}.$$

It follows from the definition of $\mathcal{F}_q$ that $\gcd(q_1, q_2, \overline{q}_1, \overline{q}_2) = 1$. Indeed, note first that q_1 and q_2 may not have a common real factor. Moreover, suppose now that $\gcd(q_1, q_2, \overline{q}_1, \overline{q}_2) = (s - \alpha)$, $\alpha \in \mathbb{C} \setminus \{0\}$. Then $(s - \overline{\alpha})$ is also a common factor of q_1 and q_2 which implies that $(s - \alpha)(s - \overline{\alpha}) \in \mathbb{R}[s]$ is a common factor of q_1 and q_2. But this is impossible according to the definition of $\mathcal{F}_q$.

By [45, Theorem 1.7], the first element of the complex Smith form of a complex polynomial matrix $R(s)$ is the greatest common divisor of all the entries of $R(s)$. Therefore, the complex Smith form of $q^c(s)$ is $\Gamma = \operatorname{diag}(\mathcal{F}_q, x)$, where

$$\mathcal{F}_q x = \det(q^c) = \mathcal{F}_q^2(q_1\overline{q}_1 + q_2\overline{q}_2), \tag{3.28}$$

and hence, noting that $\mathcal{F}_q \not\equiv 0$, $x = \mathcal{F}_q(q_1\overline{q}_1 + q_2\overline{q}_2)$. By direct calculation, $\mathcal{Q}_q\overline{\mathcal{Q}_q} = q_1\overline{q}_1 + q_2\overline{q}_2 \in \mathbb{R}[s]$. Thus, by Lemma 2.2.21-1,

$$x = \mathcal{F}_q\mathcal{Q}_q\overline{\mathcal{Q}_q} = \mathcal{M}_q$$

and the result follows. □

Combining Proposition 3.4.12 and Lemma 2.2.23 we get the following corollary.

Corollary 3.4.13. *Given a monic $q \in \mathbb{H}[s]$, there exists $p \in \mathbb{C}[s]$ such that the matrices*

$$q^c, \quad \begin{bmatrix} p & 0 \\ 0 & \overline{p} \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} \mathcal{F}_p & 0 \\ 0 & \mathcal{M}_p \end{bmatrix}$$

are equivalent.

Proof. Let $q \in \mathbb{H}[s]$. By Lemma 2.2.23 there exists $p \in \mathbb{C}[s]$ such that $\mathcal{F}_p = \mathcal{F}_q$ and $\mathcal{M}_p = \mathcal{M}_q$ and then, by Proposition 3.4.12, the matrices

$$q^c \qquad \text{and} \qquad \begin{bmatrix} \mathcal{F}_q & 0 \\ 0 & \mathcal{M}_q \end{bmatrix} = \begin{bmatrix} \mathcal{F}_p & 0 \\ 0 & \mathcal{M}_p \end{bmatrix}$$

are equivalent. Moreover, since $p \in \mathbb{C}[s]$, $p^c = \operatorname{diag}(p, \overline{p})$ and hence, again by Proposition 3.4.12, we conclude that the matrices $\operatorname{diag}(p, \overline{p})$ and $\operatorname{diag}(\mathcal{F}_p, \mathcal{M}_p)$ are also equivalent. □

In the next theorem we give the relation between the quaternionic Smith form of $R \in \mathbb{H}^{g\times r}[s]$ and the complex Smith form of R^c.

Theorem 3.4.14. *If* $\Gamma = \operatorname{diag}\left(\gamma_1, \ldots, \gamma_n\right) \in \mathbb{H}^{g\times r}[s]$ *is a quaternionic Smith form of* $R \in \mathbb{H}^{g\times r}[s]$*, then the complex Smith form of the complex adjoint matrix* $R^c \in \mathbb{C}^{2g\times 2r}[s]$ *is*

$$\Delta = \operatorname{diag}\left(\mathcal{F}_{\gamma_1}, \mathcal{M}_{\gamma_1}, \ldots, \mathcal{F}_{\gamma_n}, \mathcal{M}_{\gamma_n}\right) \in \mathbb{R}^{2g\times 2r}[s].$$

Proof. Let $R \in \mathbb{H}^{g\times r}[s]$ and suppose that $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \in \mathbb{H}^{g\times r}[s]$ is a quaternionic Smith form of R. By Proposition 3.4.12, for each $l = 1, \ldots, n$, γ_l^c is equivalent to $\operatorname{diag}\left(\mathcal{F}_{\gamma_l}, \mathcal{M}_{\gamma_l}\right)$ and consequently, Γ^c is equivalent to

$$\Delta = \operatorname{diag}\left(\mathcal{F}_{\gamma_1}, \mathcal{M}_{\gamma_1}, \ldots, \mathcal{F}_{\gamma_n}, \mathcal{M}_{\gamma_n}\right).$$

Thus, in order to show that Δ is the complex Smith form of R^c, it only remains to prove that Δ satisfies the required division properties. Obviously, $\mathcal{F}_{\gamma_l} \,|\, \mathcal{M}_{\gamma_l}$, $l = 1, \ldots, n$. Moreover, since $\gamma_l \,\|\, \gamma_{l+1}$, by Proposition 2.2.29 we have that $\mathcal{M}_{\gamma_l} \,|\, \mathcal{F}_{\gamma_{l+1}}$. Therefore Δ is indeed the complex Smith form of R^c. □

Since the complex Smith form is unique, the previous theorem allows to give an alternative proof of Corollary 3.4.9 without invoking Proposition 2.3.3.

Alternative proof of Corollary 3.4.9 Let

$$\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \quad \text{and} \quad \Gamma' = \operatorname{diag}(\gamma_1', \ldots, \gamma_n')$$

be two quaternionic Smith forms of a quaternionic matrix $R \in \mathbb{H}^{g\times r}[s]$. By Theorem 3.4.14, the complex Smith form of R^c is

$$\begin{aligned} \Delta &= \operatorname{diag}\left(\mathcal{F}_{\gamma_1}, \mathcal{M}_{\gamma_1}, \ldots, \mathcal{F}_{\gamma_n}, \mathcal{M}_{\gamma_n}\right) \\ &= \operatorname{diag}\left(\mathcal{F}_{\gamma_1'}, \mathcal{M}_{\gamma_1'}, \ldots, \mathcal{F}_{\gamma_n'}, \mathcal{M}_{\gamma_n'}\right). \end{aligned}$$

Therefore, for each $l = 1, \ldots, n$, $\mathcal{F}_{\gamma_l} = \mathcal{F}_{\gamma_l'}$ and $\mathcal{M}_{\gamma_l} = \mathcal{M}_{\gamma_l'}$, which, by Lemma 2.2.21, is equivalent to $\gamma_l\overline{\gamma}_l = \gamma_l'\overline{\gamma'}_l$. □

The following theorem characterizes the form of the complex Smith form of complex polynomial matrices with adjoint structure.

Theorem 3.4.15.

1. *If $R^c(s)$ is a complex adjoint matrix, then its complex Smith form has the structure*

$$\Delta = \operatorname{diag}(\delta_1, \delta_1', \ldots, \delta_n, \delta_n') \in \mathbb{R}^{2g\times 2r}[s], \tag{3.29}$$

where $\delta_1, \delta_1', \ldots, \delta_n, \delta_n'$ are $2n$ monic real polynomials such that $\delta_1|\delta_1'|\cdots|\delta_n|\delta_n'$ and, for every $l = 1, \ldots, n$, δ_l and δ_l' have the same real zeros with equal multiplicities, i.e., $\mu_\lambda(\delta_l) = \mu_\lambda(\delta_l')$, $\forall \lambda \in \mathbb{R}$.

2. *Consider a matrix Δ as in* (3.29). *Then there exists a complex adjoint matrix with complex Smith form Δ.*

Proof. 1. If $\Delta(s)$ is the complex Smith form of a complex adjoint matrix $R^c(s)$, by Theorem 3.4.14 there exist $\gamma_l \in \mathbb{H}[s]$, $l = 1, \ldots, n$, such that

$$\Delta = \operatorname{diag}\left(\mathcal{F}_{\gamma_1}, \mathcal{M}_{\gamma_1}, \ldots, \mathcal{F}_{\gamma_n}, \mathcal{M}_{\gamma_n}\right).$$

By definition, the polynomials $\mathcal{F}_{\gamma_l}$ and $\mathcal{M}_{\gamma_l}$ are real and, as shown in Theorem 3.4.14, $\mathcal{F}_{\gamma_1}|\mathcal{M}_{\gamma_1}|\cdots|\mathcal{F}_{\gamma_n}|\mathcal{M}_{\gamma_n}$. Moreover, by Lemma 2.2.21-1,

$$\mathcal{M}_{\gamma_l} = \mathcal{F}_{\gamma_l}\mathcal{Q}_{\gamma_l}\overline{\mathcal{Q}_{\gamma_l}},$$

and $\mathcal{Q}_{\gamma_l}$ does not have real zeros, which implies that $\mathcal{F}_{\gamma_l}$ and $\mathcal{M}_{\gamma_l}$ have the same real zeros with equal multiplicities.

2. Let $\delta_1, \delta_1', \ldots, \delta_n, \delta_n'$ be as in the theorem statement. Then, for every $l = 1, \ldots, n$, $\delta_l' = \delta_l x_l$ for some $x_l \in \mathbb{R}[s]$ with no real zeros. This implies that the zeros of x_l appears in conjugate pairs and therefore $x_l = c_l\bar{c}_l$, for some $c_l \in \mathbb{C}[s]$ with no real zeros and then $\delta_l' = \delta_l c_l \bar{c}_l$. Let $p_l = \delta_l c_l \in \mathbb{C}[s]$. By definition, $\mathcal{F}_{p_l} = \delta_l$ and $\mathcal{Q}_{p_l} = c_l$ and therefore $\mathcal{M}_{p_l} = \delta_l c_l \bar{c}_l$. Thus, by Corollary 3.4.13, $\operatorname{diag}(\delta_l, \delta_l') = \operatorname{diag}\left(\mathcal{F}_{p_l}, \mathcal{M}_{p_l}\right)$ is equivalent to $\operatorname{diag}(p_l, \bar{p}_l)$. Hence, the matrix Δ is equivalent to

$$\operatorname{diag}(p_1, \overline{p_1}, \ldots, p_n, \overline{p_n}),$$

which, in turn, is equivalent to the complex adjoint matrix, R^c, of

$$R = \mathrm{diag}(p_1, \ldots, p_n) \in \mathbb{H}^{g\times r}[s].$$

□

3.5 Quaternionic Smith-McMillan form

As we shall see in Section 4.2.3, quaternionic rational matrices play an important role in the study of input/output quaternionic systems. In this section we define such matrices and analyze their algebraic properties. In particular, we study the quaternionic Smith-McMillan form since it turns out that, as happens in the commutative case, this form is relevant in the characterization of the BIBO-stability. Moreover, we characterize the complex Smith-McMillan form of complex adjoint matrices and we give the relation between the quaternionic Smith-McMillan of a quaternionic matrix and the complex Smith-McMillan form of its complex adjoint matrix.

We start by recalling the notion of Smith-McMillan form of real and complex rational matrices.

Theorem 3.5.1. [25, 45] *Let $R \in \mathbb{R}^{g\times r}(s)$ be a rational matrix with rank n. Then there exist unimodular polynomial matrices U and V such that*

$$URV = \mathrm{diag}\left(\frac{\epsilon_1}{\psi_1}, \ldots, \frac{\epsilon_n}{\psi_n}\right) \in \mathbb{R}^{g\times r}(s), \tag{3.30}$$

where ϵ_l and ψ_l are monic polynomials that satisfy the following conditions:

- *the fraction $\frac{\epsilon_l}{\psi_l}$ is irreducible, i.e., (ϵ_l, ψ_l) are coprime;*
- $\epsilon_l \,|\, \epsilon_{l+1}$ *and* $\psi_{l+1} \,|\, \psi_l$.

The matrix (3.30) is the *complex Smith-McMillan form* of R.

Before extending the previous theorem to the quaternionic case, we need to give a formal definition of rational functions with coefficients in the quaternionic skew-field.

For that purpose we first introduce the concept of *Ore ring* [38] and prove that $\mathbb{H}[s]$ is such a ring.

Definition 3.5.2. [38] An integrity domain $\mathcal{R}$ is said to be a *Ore ring* if any two nonzero elements $a, b \in \mathcal{R}$ have a left and a right nonzero common multiple, i.e., for instance on the left, there exist $\alpha, \beta \in \mathcal{R}$ such that $\alpha a = \beta b \neq 0$.

Proposition 3.5.3. [37, 38] *The division ring* $\mathbb{H}[s]$ *is an Ore ring.*

Proof. Let $p(s)$ and $q(s)$ be nonzero quaternionic polynomials. We will show by induction that they have a right common multiple. Suppose that $q(s)$ has degree zero, i.e., $q(s) = c \in \mathbb{H}\backslash\{0\}$. Then

$$p(s) \cdot 1 = q(s)\big(c^{-1}p(s)\big),$$

i.e., $p(s)$ is a right common multiple of $p(s)$ and $q(s)$. Since the roles of p and q are interchangeable, this shows the existence of a common right multiple in case one of the polynomials has degree zero.

Let $n \geq 1$ and suppose now that a common right multiple exists in case one of the polynomials has degree not higher than $n-1$. We will see that this implies that the property still holds if one of the polynomials has degree n. Assume then that, for instance, $\deg q = n$ and $\deg p \geq n$ (in case $\deg p < n$ the result follows from the induction hypothesis). Applying the Euclidian algorithm we have that

$$p(s) = q(s)d(s) + r(s),$$

where $r(s)$ is the zero polynomial or has degree strictly lower than n. If $r(s) = 0$ then

$$p(s) \cdot 1 = q(s)d(s).$$

Otherwise, by the induction hypothesis, there exist nonzero polynomials $u(s)$ and $v(s)$ such that

$$q(s)u(s) = r(s)v(s)$$

and therefore

$$p(s)v(s) = q(s)\big[d(s)v(s) + u(s)\big].$$

The proof that $p(s)$ and $q(s)$ have a left common multiple is similar. Therefore we can conclude that $\mathbb{H}[s]$ is an Ore ring. $\square$

Since $\mathbb{H}[s]$ is an Ore ring it is possible to construct its field of fractions [37, 38]. Let $p, p' \in \mathbb{H}[s]$, $q, q' \in \mathbb{H}[s] \setminus \{0\}$ and consider the fractions $\frac{p}{q}$ and $\frac{p'}{q'}$. We say that

$$\frac{p}{q} = \frac{p'}{q'}$$

if for every $\alpha, \beta \in \mathbb{H}[s]$ such that $\alpha q = \beta q'$, then $\alpha p = \beta p'$. Note that, with this definition, $\frac{p}{q} = \frac{dp}{dq}$, $dq \neq 0$, which is the simplification rule on the left. Therefore, we say that the fraction $\frac{p}{q}$ is irreducible if and only if the pair (p, q) is left coprime, i.e., if p and q only have trivial (constant) common left divisors.

Note that it is not possible to apply the simplification rule on the right. For example, $\frac{\mathbf{i}}{1} \neq \frac{\mathbf{ij}}{\mathbf{j}} = \frac{\mathbf{k}}{\mathbf{j}}$. Indeed, they are equal if and only if

$$\alpha 1 = \beta \mathbf{j} \;\Rightarrow\; \alpha \mathbf{i} = \beta \mathbf{k}$$

for every $\alpha, \beta \in \mathbb{H}[s]$. But, if $\alpha = \mathbf{j}$ and $\beta = 1$, the hypothesis is satisfied while the thesis is $\alpha \mathbf{i} = \mathbf{ji} = -\mathbf{k} \neq \beta \mathbf{k} = \mathbf{k}$.

The set of left quaternionic rational functions is defined as

$$\mathbb{H}(s) = \left\{ \frac{p}{q} : \; p, q \in \mathbb{H}[s], \; q \neq 0 \right\}.$$

Given two fractions $\frac{p}{q}$ and $\frac{p'}{q'} \in \mathbb{H}(s)$, their addition and product are defined, respectively, as follows:

$$\frac{p}{q} + \frac{p'}{q'} = \frac{\alpha p + \beta p'}{\beta q'}$$

where $\alpha q = \beta q'$, $\alpha, \beta \in \mathbb{H}[s] \setminus \{0\}$ and

$$\frac{p}{q} \cdot \frac{p'}{q'} = \frac{\delta p'}{\gamma q}$$

where $\delta q' = \gamma p$, $\delta, \gamma \in \mathbb{H}[s]$, $\gamma \neq 0$.

It follows from the product definition that the left inverse of $\frac{p}{q}$, $p, q \neq 0$ is $\frac{q}{p}$ and this proves that $\mathbb{H}(s)$ is a skew-field. We will identify $\mathbb{H}[s] \cong \left\{ \frac{q}{1}, \; q \in \mathbb{H}[s] \right\}$ and thus denote by q^{-1} the left inverse of $q \in \mathbb{H}[s]$, i.e., $q^{-1} = \frac{1}{q}$. Therefore we shall write $q^{-1}p = \frac{p}{q}$ since $\frac{1}{q} \cdot \frac{p}{1} = \frac{p}{q}$.

The definition of $\mathbb{H}(s, s^{-1})$ is analogous. As usual, $\mathbb{H}^{g\times r}(s)$ and $\mathbb{H}^{g\times r}(s, s^{-1})$ will denote, respectively, the set of the $g \times r$ matrices with entries in $\mathbb{H}(s)$ and $\mathbb{H}(s, s^{-1})$.

The definition of the complex adjoint matrix of $R(s) \in \mathbb{H}^{g\times r}(s)$, $R^c(s)$, is analogous to the polynomial case and Theorem 3.3.1 also holds for rational quaternionic matrices.

We are now in a position to define the Smith-McMillan form of quaternionic rational matrices.

Theorem 3.5.4. *Let $R \in \mathbb{H}^{g\times r}(s)$ be a rational quaternionic matrix with rank n. Then there exist quaternionic unimodular polynomial matrices U and V such that*

$$URV = \operatorname{diag}\left(\frac{\epsilon_1}{\psi_1}, \dots, \frac{\epsilon_n}{\psi_n}\right) \in \mathbb{H}^{g\times r}(s), \tag{3.31}$$

where ϵ_l and ψ_l are monic polynomials that satisfy the following conditions for any l:

- *the fraction $\frac{\epsilon_l}{\psi_l}$ is irreducible, i.e., (ϵ_l, ψ_l) are left coprime;*
- *$\epsilon_l \parallel \epsilon_{l+1}$ and $\psi_{l+1} \parallel \psi_l$.*

The matrix (3.31) is a *quaternionic Smith-McMillan form* of R.

Proof. Let $d \in \mathbb{R}[s]$ be monic and such that $M = dR \in \mathbb{H}^{g\times r}[s]$ is a polynomial matrix. Such a polynomial always exists (for instance, let d be the least common real multiple of the denominators of the entries of R). By Theorem 3.4.3, M can be reduced to a quaternionic Smith form, $\operatorname{diag}(\gamma_1, \dots, \gamma_n)$ where $\gamma_l \parallel \gamma_{l+1}$, by means of unimodular polynomial matrices U and V such that

$$UMV = \operatorname{diag}(\gamma_1, \dots, \gamma_n). \tag{3.32}$$

Since d is a real polynomial it commutes with the matrix U, and hence, dividing both sides of (3.32) by d, we obtain

$$URV = Ud^{-1}MV = d^{-1}UMV = \operatorname{diag}\left(\frac{\gamma_1}{d}, \dots, \frac{\gamma_n}{d}\right).$$

Therefore, by eliminating the common left factors of the fractions $\frac{\gamma_l}{d}$, we obtain the matrix (3.31) with irreducible fractions. It only remains to show that the numerators and the denominators verify the required properties.

Let $\alpha_l \in \mathbb{H}[s]$ be the greatest common monic left factor of γ_l and of d. Then we can write $\gamma_l = \alpha_l \epsilon_l$ and $d = \alpha_l \psi_l$. Because $d \in \mathbb{R}[s]$, by Proposition 2.2.30-2, we obtain $\mathcal{Q}_{\psi_l} = \overline{\mathcal{Q}_{\alpha_l}}$. Moreover, since ϵ_l and ψ_l are left coprime, also $\mathcal{Q}_{\epsilon_l}$ and $\mathcal{Q}_{\psi_l} = \overline{\mathcal{Q}_{\alpha_l}}$ are left coprime and so, by Proposition 2.2.30-1

$$\mathcal{M}_{\gamma_l} = \mathcal{M}_{\alpha_l \epsilon_l} = \mathcal{M}_{\alpha_l}\mathcal{M}_{\epsilon_l}.$$

Taking into account that $\gamma_l \parallel \gamma_{l+1}$, it follows from Proposition 2.2.29 that $\mathcal{M}_{\gamma_l} \mid \gamma_{l+1}$, i.e., $\gamma_{l+1} = \mathcal{M}_{\gamma_l}\beta$ for some $\beta \in \mathbb{H}[s]$. Further, since $d = \alpha_l \psi_l \in \mathbb{R}[s]$, by Corollary 2.2.31, $d = \mathcal{M}_{\alpha_l}\mathcal{F}_{\psi_l}$. Therefore,

$$\frac{\gamma_{l+1}}{d} = \frac{\mathcal{M}_{\alpha_l}\mathcal{M}_{\epsilon_l}\beta}{\mathcal{M}_{\alpha_l}\mathcal{F}_{\psi_l}} = \frac{\mathcal{M}_{\epsilon_l}\beta}{\mathcal{F}_{\psi_l}} = \frac{\epsilon_{l+1}}{\psi_{l+1}}.$$

Note that in the last equality, simplifications can only occur between β and $\mathcal{F}_{\psi_l}$ since $\mathcal{M}_{\epsilon_l}$ and $\mathcal{F}_{\psi_l}$ are left coprime due to the left coprimeness of ϵ_l and ψ_l. This clearly shows that $\mathcal{M}_{\epsilon_l} | \epsilon_{l+1}$ and that $\psi_{l+1} | \mathcal{F}_{\psi_l}$, i.e., by Proposition 2.2.29, that $\epsilon_l \parallel \epsilon_{l+1}$ and $\psi_{l+1} \parallel \psi_l$. $\square$

The proof of the previous theorem shows the following deep relation between quaternionic Smith forms and quaternionic Smith-McMillan forms.

Corollary 3.5.5. *Let $R \in \mathbb{H}^{g\times r}(s)$ be a rational quaternionic matrix and write $R = d^{-1}M$, where $d \in \mathbb{R}[s]$ is the least common real multiple of the denominators of the entries of R and $M \in \mathbb{H}^{g\times r}[s]$. Then Ψ is a quaternionic Smith-McMillan form of R if and only if $d\Psi$ is a quaternionic Smith form of M.*

Remark 3.5.6. The polynomials ϵ_l and ψ_l in (3.31) may have common zeros. For example,

$$\frac{\epsilon(s)}{\psi(s)} = \frac{\mathbf{j}(s-\mathbf{i})}{s-\mathbf{i}}$$

is a quaternionic Smith-McMillan form. Actually, ϵ_l and ψ_l are left coprime since they satisfy the Bezout equation

$$\epsilon(s) \cdot \left(-\tfrac{\mathbf{k}}{2}\right) + \psi(s) \cdot \tfrac{\mathbf{i}}{2} = 1.$$

However, $\epsilon(\mathbf{i}) = \psi(\mathbf{i}) = 0$, which implies that the polynomials ϵ_l and ψ_l have a common zero, or in other words, are not zero coprime. □

As happens with the quaternionic Smith form, the quaternionic Smith-McMillan form is not unique.

Example 3.5.7. Consider the quaternionic rational matrix

$$R(s) = \begin{bmatrix} \frac{1}{s^2+1} & 0 \\ 0 & \frac{s+\mathbf{i}}{s^2+1} \end{bmatrix}$$

and let

$$M(s) = (s^2+1)R(s) = \begin{bmatrix} 1 & 0 \\ 0 & s+\mathbf{i} \end{bmatrix}.$$

By Example 3.4.5, the matrices

$$\gamma(s) = \begin{bmatrix} 1 & 0 \\ 0 & s+\mathbf{i} \end{bmatrix} \qquad \text{and} \qquad \gamma'(s) = \begin{bmatrix} 1 & 0 \\ 0 & s-\mathbf{i} \end{bmatrix}$$

are two quaternionic Smith forms of M. Therefore, the matrices with irreducible fractions

$$(s^2+1)^{-1}\gamma(s) = \begin{bmatrix} \frac{1}{s^2+1} & 0 \\ 0 & \frac{1}{s-\mathbf{i}} \end{bmatrix} \quad \text{and} \quad (s^2+1)^{-1}\gamma'(s) = \begin{bmatrix} \frac{1}{s^2+1} & 0 \\ 0 & \frac{1}{s+\mathbf{i}} \end{bmatrix}$$

are two different quaternionic Smith-McMillan forms of the matrix R. □

Analogously to what happens with the quaternionic Smith form, the source of nonuniqueness of the Smith-McMillan form can be characterized in the following way.

Lemma 3.5.8. *Let $R \in \mathbb{H}^{g\times r}(s)$ be a rational quaternionic matrix and $d \in \mathbb{R}[s]$ be the least common real multiple of the denominators of the entries of R. Suppose that*

$$\Psi = \operatorname{diag}\left(\frac{\epsilon_1}{\psi_1}, \ldots, \frac{\epsilon_n}{\psi_n}\right) \in \mathbb{H}^{g\times r}(s)$$

is a quaternionic Smith-McMillan form of R.

1. *If*
$$\Psi' = \operatorname{diag}\left(\frac{\epsilon_1'}{\psi_1'}, \ldots, \frac{\epsilon_n'}{\psi_n'}\right) \in \mathbb{H}^{g\times r}(s)$$
is another quaternionic Smith-McMillan form of R *then*
$$d\frac{\epsilon_l}{\psi_l} \sim_J d\frac{\epsilon_l'}{\psi_l'}, \quad l = 1, \ldots n.$$

2. *If* $\operatorname{rank} R(s) \geq 2$ *and* $d\frac{\epsilon_l}{\psi_l} \sim_J d\frac{\epsilon_l'}{\psi_l'}$, $l = 1, \ldots n$, *then*
$$\Psi' = \operatorname{diag}\left(\frac{\epsilon_1'}{\psi_1'}, \ldots, \frac{\epsilon_n'}{\psi_n'}\right) \in \mathbb{H}^{g\times r}(s)$$
is also a quaternionic Smith-McMillan form of R.

Proof. Note that, by the definition of $d \in \mathbb{R}[s]$, $\psi_l \mid d$ and $\psi_l' \mid d$, $l = 1, \ldots, n$. Therefore $d\frac{\epsilon_l}{\psi_l}$ and $d\frac{\epsilon_l'}{\psi_l'}$ are polynomials, i.e., they are elements of $\mathbb{H}[s]$. The first point follows from Corollary 3.5.5 and Theorem 3.4.8 while se second follows by Corollary 3.5.5 and Theorem 3.4.10. □

The following theorem gives the relation between a quaternionic Smith-McMillan form of a quaternionic rational matrix and the complex Smith-McMillan form of its complex adjoint matrix. This relation is an extension of the one given in Theorem 3.4.14 for Smith forms.

Theorem 3.5.9. *Let* $R \in \mathbb{H}^{g\times r}(s)$ *and*
$$\operatorname{diag}\left(\frac{\epsilon_1}{\psi_1}, \ldots, \frac{\epsilon_n}{\psi_n}\right)$$
be a quaternionic Smith-McMillan form of R. *Then the complex Smith-McMillan form of the corresponding complex adjoint matrix* R^c *is given by*
$$\operatorname{diag}\left(\frac{\mathcal{F}_{\epsilon_1}}{\mathcal{M}_{\psi_1}}, \frac{\mathcal{M}_{\epsilon_1}}{\mathcal{F}_{\psi_1}}, \ldots, \frac{\mathcal{F}_{\epsilon_n}}{\mathcal{M}_{\psi_n}}, \frac{\mathcal{M}_{\epsilon_n}}{\mathcal{F}_{\psi_n}}\right) \in \mathbb{R}^{2g\times 2r}(s). \tag{3.33}$$

Proof. With the notation used in the proof of Theorem 3.5.4, we let
$$\frac{\epsilon_l}{\psi_l} = \frac{\alpha_l \epsilon_l}{\alpha_l \psi_l} = \frac{\gamma_l}{d}$$

where $\operatorname{diag}(\gamma_1, \ldots, \gamma_n)$ is a quaternionic Smith form, $\overline{\mathcal{Q}_{\alpha_l}}$ and $\mathcal{Q}_{\epsilon_l}$ are left coprime, and d a real monic polynomial. By Proposition 2.2.30, it follows that

$$\mathcal{M}_{\gamma_l} = \mathcal{M}_{\alpha_l}\mathcal{M}_{\epsilon_l} \quad \text{and} \quad \mathcal{F}_{\gamma_l} = \mathcal{F}_{\alpha_l}\mathcal{F}_{\epsilon_l}.$$

Moreover, by Corollary 2.2.31, $d = \mathcal{M}_{\alpha_l}\mathcal{F}_{\psi_l} = \mathcal{F}_{\alpha_l}\mathcal{M}_{\psi_l}$. Therefore, applying Theorem 3.4.14 the result follows. □

Theorem 3.4.15 states that the complex Smith form of any complex adjoint matrix has a special structure. The following theorem shows that the same holds true for Smith-McMillan forms.

Theorem 3.5.10.

1. *If $R^c(s) \in \mathbb{C}^{2g\times 2r}(s)$ is a complex adjoint matrix, then its complex Smith-McMillan form has the structure*

$$\Omega = \operatorname{diag}\left(\frac{\theta_1}{\omega_1}, \frac{\theta_1'}{\omega_1'}, \ldots, \frac{\theta_n}{\omega_n}, \frac{\theta_n'}{\omega_n'}\right) \in \mathbb{R}^{2g\times 2r}(s), \tag{3.34}$$

   where $\theta_l, \theta_l', \omega_l, \omega_l'$, $l = 1, \ldots, n$, are $4n$ monic real polynomials such that the fractions $\frac{\theta_l}{\omega_l}$ and $\frac{\theta_l'}{\omega_l'}$ are irreducible, $\theta_1|\theta_1'|\cdots|\theta_n|\theta_n'$, $\omega_n'|\omega_n|\cdots|\omega_1'|\omega_1$, $\mu_\lambda(\theta_l) = \mu_\lambda(\theta_l')$ and $\mu_\lambda(\omega_l) = \mu_\lambda(\omega_l')$, $\forall \lambda \in \mathbb{R}$.

2. *Consider a matrix Ω as in* (3.34)*. Then there exists a complex adjoint matrix with complex Smith-McMillan form Ω.*

Proof. 1. This is a consequence of Theorem 3.5.9.

2. Let Ω be a matrix as in (3.34). Let $d \in \mathbb{R}[s]$ be the least common multiple of the denominators ω_l, ω_l' of (3.34) and define d_l, d_l' by the relations $d_l\omega_l = d_l'\omega_l' = d$ for $l = 1, \ldots, n$. Note that, for every $\lambda \in \mathbb{R}$,

$$\mu_\lambda(d) = \mu_\lambda(d_l) + \mu_\lambda(\omega_l) = \mu_\lambda(d_l') + \mu_\lambda(\omega_l').$$

Therefore the condition on the zeros of ω_l and of ω_l' is consequently satisfied also by d_l and by d_l' and so, by hypothesis, it also holds true for $\delta_l = d_l\theta_l$ and $\delta_l' = d_l'\theta_l'$. By

Theorem 3.4.15,

$$\mathrm{diag}(\delta_1, \delta_1', \ldots, \delta_n, \delta_n') \in \mathbb{R}^{2g\times 2r}[s]$$

is the complex Smith form of M^c, for some $M \in \mathbb{H}^{g\times r}[s]$, which clearly implies that (3.34) is the complex Smith-McMillan form of the complex adjoint matrix of $d^{-1}M$. $\square$

Chapter 4

Quaternionic behavioral systems

In this chapter we extend the behavioral approach introduced by Willems [53, 54] to quaternionic systems. We start by studying behaviors that can be described as solution sets of quaternionic matrix difference or differential equations, i.e., those which are the kernel of some suitable matrix difference or differential operator. Difference equations arise either directly or from the digital implementation of problems described by differential equations. We will use the terminology dynamical equations to refer to both difference or differential equations, depending on the situation.

Besides kernel representations, we also analyze other descriptions of a system, such as image and input-output representations. Finally we give a complete and explicit characterization of all solutions of quaternionic matrix difference and differential equations.

4.1 Basic definitions

The main difference between the classical and the behavioral approaches is that while the second starts by dividing the system variables into input, output and/or variables, the behavioral approach looks at the system variables without imposing a priori any structure, i.e., without speaking, at an early stage, of inputs and outputs, of causes and effects.

In the sequel we will give some basic definitions concerning behavioral systems, [43].

Definition 4.1.1. A *dynamical system* Σ is defined as a triple of sets

$$\Sigma = (\mathbb{T}, \mathbb{W}, \mathcal{B}),$$

where $\mathbb{T}$ is called the *time axis*, $\mathbb{W}$ is the *signal space*, and $\mathcal{B}$ is a subset of $\mathbb{W}^{\mathbb{T}} = \{f : \mathbb{T} \to \mathbb{W}\}$ called the system *behavior*. The elements of $\mathcal{B}$ are called *trajectories*.

In this thesis $\mathbb{T} = \mathbb{Z}$ or $\mathbb{T} = \mathbb{R}$, and $\mathbb{W} = \mathbb{H}^r$, for some $r \in \mathbb{N}$, i.e., we consider *discrete* or *continuous-time quaternionic systems*, respectively. In the sequel we define fundamental properties of behaviors.

Definition 4.1.2. A behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$ is said to be *linear on the right* [*left*] if for any $w_1, w_2 \in \mathcal{B}$ and $\alpha_1, \alpha_2 \in \mathbb{H}$, $w_1\alpha_1 + w_2\alpha_2 \in \mathcal{B}$ $[\alpha_1 w_1 + \alpha_2 w_2 \in \mathcal{B}]$. A behavior is said to be *linear* if it is both linear on the right and on the left.

If $\tau \in \mathbb{T}$ is such that $t + \tau \in \mathbb{T}$ for every $t \in \mathbb{T}$, we define the *τ-shift operator* by $(\sigma^\tau w)(t) = w(t + \tau)$, for $t \in \mathbb{T}$.

Definition 4.1.3. A behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$, with $\mathbb{T} = \mathbb{Z}$ or $\mathbb{T} = \mathbb{R}$, is said to be *shift-invariant* if for every $w \in \mathcal{B}$, $\sigma^\tau w \in \mathcal{B}$, $\forall \tau \in \mathbb{T}$, or, equivalently, $\sigma^\tau \mathcal{B} \subset \mathcal{B}$, $\forall \tau \in \mathbb{T}$.

Consider a behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$. We denote by $w|_I$ the restriction of a trajectory $w \in \mathcal{B}$ to a certain time interval $I \subset \mathbb{T}$ and by $\mathcal{B}|_I$ the set of all these trajectories, i.e.,

$$\mathcal{B}|_I = \{w|_I \ : \ w \in \mathcal{B}\}.$$

Definition 4.1.4. A behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$ is said to be *complete* if

$$\{w \in \mathcal{B}\} \Leftrightarrow \{w|_I \in \mathcal{B}|_I, \text{ for all finite intervals } I \subset \mathbb{T}\}.$$

This means that, when a behavior is complete, we just need to check locally whether a trajectory satisfies or not the system laws.

4.2 Behavior representations

In this section we present some representations of quaternionic behaviors. We start with the most general ones, the kernel representations, and characterize the equivalence between two such representations of the same behavior. Then other representations such as image and input-output representations are introduced.

4.2.1 Kernel representations

The behavior $\mathcal{B}$ of a system can often be described by means of a linear, homogeneous matrix difference or differential equation with constant coefficients, i.e., the trajectories w in $\mathcal{B}$ are the solutions of an equation of the form

$$\sum_{l=M}^{N} R_l w(t+l) = 0, \qquad \forall\, t \in \mathbb{Z}, \tag{4.1}$$

with $M, N \in \mathbb{Z}$, in the discrete case, or,

$$\sum_{l=0}^{N} R_l w^{(l)}(t) = 0, \qquad \forall\, t \in \mathbb{R}, \tag{4.2}$$

with $N \in \mathbb{Z}_0^+$, in the continuous case, where $R_l \in \mathbb{H}^{g\times r}$ are constant matrices. In the continuous case, $w^{(l)}$ is the l-th order derivative of w and trajectories are supposed to be sufficiently smooth or otherwise equations are to be understood in a distributional sense (see [43]). For simplicity we will assume henceforth that all admissible trajectories are $\mathcal{C}^\infty$.

Using the shift operator σ introduced in the previous section we have that $w(t+l) = \sigma^l w(t)$ and equation (4.1) can be written in the more compact form

$$R(\sigma, \sigma^{-1})w(t) = 0, \tag{4.3}$$

where $R(\sigma, \sigma^{-1}) = \sum_{l=M}^{N} R_l \sigma^l$. This amounts to say that $\mathcal{B}$ is the kernel of the difference operator $R(\sigma, \sigma^{-1})$ acting on $(\mathbb{H}^r)^{\mathbb{Z}}$, i.e.,

$$\mathcal{B} = \ker R(\sigma, \sigma^{-1}) = \left\{ w \in (\mathbb{H}^r)^{\mathbb{Z}} : R(\sigma, \sigma^{-1})w = 0 \right\}. \tag{4.4}$$

The form of the operator $R(\sigma,\sigma^{-1})$ in (4.3) suggests, as it is usual within the behavioral approach, to consider the polynomial matrix in s and s^{-1}

$$R(s,s^{-1}) = \sum_{l=M}^{N} R_l s^l,$$

which is called a *kernel representation* of the behavior (4.4).

In order to make the composition of operators correspond to the multiplication of the corresponding matrices, we view $R(s,s^{-1})$ as an element of $\mathbb{H}^{g\times r}[s,s^{-1}]$, cf Section 2.1.

When dealing with continuous systems we consider matrices with entries in $\mathbb{H}[s]$. If $R(s) = \sum_{l=0}^{N} R_l s^l \in \mathbb{H}^{g\times r}[s]$, equation (4.2) can be written in the operator form as

$$R\left(\frac{d}{dt}\right) w(t) = \sum_{l=0}^{N} R_l \frac{d^l}{dt^l} w(t) = 0.$$

Thus

$$\mathcal{B} = \ker R\left(\frac{d}{dt}\right) = \left\{ w \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{H}^r) : R\left(\frac{d}{dt}\right) w = 0 \right\}.$$

Both in the discrete and in the continuous case, we refer to behaviors with kernel representations as *kernel behaviors*. Throughout this thesis we will often simply denote kernel behaviors as $\mathcal{B} = \ker R$, both in the discrete and in the continuous case, meaning that $\mathcal{B} = \ker R(\sigma,\sigma^{-1})$ when $\mathbb{T} = \mathbb{Z}$ and $\mathcal{B} = \ker R\left(\frac{d}{dt}\right)$ when $\mathbb{T} = \mathbb{R}$.

Example 4.2.1. Consider the equations $(\sigma - \alpha)^n w(t) = 0$ and $\left(\frac{d}{dt} - \alpha\right)^n w(t) = 0$ where $\alpha \in \mathbb{H}$. Their solutions are the kernel of operators represented by the polynomial $p(s) = (s-\alpha)^n$. It is not difficult to check that the solutions are $w(t) = t^l \alpha^t q$ and $w(t) = t^l e^{\alpha t} q$, respectively, for every $l = 0, \ldots, n-1$ and $q \in \mathbb{H}$. $\square$

Note that behaviors which can be written as the kernel of some difference or differential operator are shift-invariant, complete and linear on the right, but not on the left.

Example 4.2.2. Consider the quaternionic polynomial $p(s) = s - \mathbf{j}$ and let $\mathcal{B} = \ker p(\sigma)$. It is clear that $w(t) = \mathbf{j}^t \in \mathcal{B}$. However the trajectory $\mathbf{i}w(t)$ does not belong to $\mathcal{B}$. In fact,

$$(\sigma - \mathbf{j})\mathbf{i}\mathbf{j}^t = \mathbf{i}\sigma\mathbf{j}^t - \mathbf{j}\mathbf{i}\mathbf{j}^t = \mathbf{i}\mathbf{j}^{t+1} + \mathbf{k}\mathbf{j}^t = \mathbf{i}\mathbf{j}\mathbf{j}^t + \mathbf{k}\mathbf{j}^t = 2\mathbf{k}\mathbf{j}^t \neq 0$$

□

Similar to what happens in the real or complex case, a linear on the right, shift-invariant and complete discrete behavior admits a kernel representation, but the same does not necessary hold in the continuous case. In this thesis we restrict our attention to systems with kernel behaviors and we will refer to them simply as *behaviors*.

Remark 4.2.3. Clearly, quaternionic kernel behaviors are vector spaces, when equipped with the usual addition and right multiplication by scalars. □

It was shown in Section 1.2.1 that each quaternionic matrix can be related with a complex one, the complex adjoint matrix. This was generalized for polynomial matrices in Section 3.3. Similarly, we extend the map (1.9) to sequences and define for any behavior $\mathcal{B}$ the complex behavior

$$\mathcal{B}^{\mathbb{C}} = \left\{ w^{\mathbb{C}} : \, w \in \mathcal{B} \right\},$$

where $w^{\mathbb{C}}(t) = \big(w(t)\big)^{\mathbb{C}}$, $\forall\, t \in \mathbb{T}$, which we call the *complex form* of $\mathcal{B}$.

Quaternionic behaviors share many properties with their complex form as can be seen in the following lemma.

Lemma 4.2.4. *Consider a behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$ and let $\mathcal{B}^{\mathbb{C}}$ be its complex form.*

1. *If $\mathcal{B}$ is linear on the right then $\mathcal{B}^{\mathbb{C}}$ is linear.*

2. *$\mathcal{B}$ is shift-invariant if and only if $\mathcal{B}^{\mathbb{C}}$ is shift-invariant.*

3. *$\mathcal{B}$ is complete if and only if $\mathcal{B}^{\mathbb{C}}$ is complete.*

Proof. 1. Assume that $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{T}}$ is linear on the right and let $x = x_1 + x_2\mathbf{j} \in \mathcal{B}$ and $y = y_1 + y_2\mathbf{j} \in \mathcal{B}$. Then, by definition, $x^{\mathbb{C}}$ and $y^{\mathbb{C}}$ are trajectories of $\mathcal{B}^{\mathbb{C}}$. We claim that for any $\alpha, \beta \in \mathbb{C}$, $x^{\mathbb{C}}\alpha + y^{\mathbb{C}}\beta \in \mathcal{B}^{\mathbb{C}}$. Since $\mathcal{B}$ is linear on the right, $x\alpha + y\beta \in \mathcal{B}$ and then $(x\alpha + y\beta)^{\mathbb{C}} \in \mathcal{B}^{\mathbb{C}}$. Now it just remains to prove that $(x\alpha + y\beta)^{\mathbb{C}} = x^{\mathbb{C}}\alpha + y^{\mathbb{C}}\beta$. Recalling that, as mentioned in (1.2), $\mathbf{j}z = \overline{z}\mathbf{j}$ for all $z \in \mathbb{C}$,

$$x\alpha = (x_1 + x_2\mathbf{j})\alpha = x_1\alpha + x_2\overline{\alpha}\mathbf{j} \quad \text{and} \quad y\beta = (y_1 + y_2\mathbf{j})\beta = y_1\beta + y_2\overline{\beta}\mathbf{j}$$

and then

$$x\alpha + y\beta = (x_1\alpha + y_1\beta) + (x_2\overline{\alpha} + y_2\overline{\beta})\mathbf{j}.$$

Thus

$$(x\alpha + y\beta)^{\mathbb{C}} = \begin{bmatrix} x_1\alpha + y_1\beta \\ \overline{-(x_2\overline{\alpha} + y_2\overline{\beta})} \end{bmatrix} = \begin{bmatrix} x_1\alpha + y_1\beta \\ -\overline{x}_2\alpha - \overline{y}_2\beta \end{bmatrix} = \begin{bmatrix} x_1 \\ -\overline{x}_2 \end{bmatrix}\alpha + \begin{bmatrix} y_1 \\ -\overline{y}_2 \end{bmatrix}\beta = x^{\mathbb{C}}\alpha + y^{\mathbb{C}}\beta.$$

2. Let $\mathcal{B}$ be shift-invariant and $w^{\mathbb{C}}(t) \in \mathcal{B}^{\mathbb{C}}$. Then, $w(t) \in \mathcal{B}$ and, by the shift invariance, $w(t+\tau) \in \mathcal{B}$, for every $\tau \in \mathbb{T}$. But this implies that $w^{\mathbb{C}}(t+\tau) \in \mathcal{B}^{\mathbb{C}}$ and therefore $\mathcal{B}^{\mathbb{C}}$ is shift-invariant. The reciprocal is analogous.

3. The proof is similar to the one of the second point. □

Note that the reciprocal of Lemma 4.2.4-1 does not hold.

Example 4.2.5. Let $\mathcal{B} = \left\{w \in (\mathbb{H}^r)^{\mathbb{T}} : \ w = w_1 - 2\overline{w}_1\mathbf{j},\ w_1 \in (\mathbb{C}^r)^{\mathbb{T}}\right\}$. By definition,

$$\mathcal{B}^{\mathbb{C}} = \left\{ w^{\mathbb{C}} : \ w^{\mathbb{C}} = \begin{bmatrix} w_1 \\ 2w_1 \end{bmatrix} \in (\mathbb{C}^{2r})^{\mathbb{T}} \right\}.$$

It is easy to check that $\mathcal{B}^{\mathbb{C}}$ is linear. However, while $w = w_1 - 2\overline{w}_1\mathbf{j} \in \mathcal{B}$, $w\mathbf{j} = 2\overline{w}_1 + w_1\mathbf{j} \notin \mathcal{B}$ and hence $\mathcal{B}$ is not linear on the right. □

In the sequel, throughout the whole thesis, we will only consider discrete-time systems. The results for continuous-time systems are analogous.

As the following proposition shows, if $\mathcal{B}$ is a kernel behavior then so is $\mathcal{B}^c$. Moreover, its kernel representations can be derived from any kernel representation of $\mathcal{B}$.

Proposition 4.2.6. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$. Then $\left(\ker R\,(\sigma, \sigma^{-1})\,\right)^{\mathbb{C}} = \ker R^c\,(\sigma, \sigma^{-1})$.*

Proof. Let $v \in (\ker R)^{\mathbb{C}}$. Then, by definition there exists $w \in \ker R$ such that $v = w^{\mathbb{C}}$. Since $Rw = 0$ then $R^c v = R^c w^{\mathbb{C}} = 0$. Hence $v \in \ker R^c$. Conversely, let $v \in \ker R^c$. This uniquely determines w such that $v = w^{\mathbb{C}}$. Then $(Rw)^{\mathbb{C}} = R^c w^{\mathbb{C}} = R^c v = 0$, which implies that $w \in \ker R$ and, since $v = w^{\mathbb{C}}$, $v \in (\ker R)^{\mathbb{C}}$. □

On the other hand, if $\mathcal{B}^{\mathbb{C}} = \ker R$ then there exists a quaternionic matrix $\widetilde{R}$ such that $\mathcal{B} = \ker \widetilde{R}$.

Proposition 4.2.7. *Consider a behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ linear on the right and let $\mathcal{B}^{\mathbb{C}} = \ker R\,(\sigma, \sigma^{-1})$, $R \in \mathbb{C}^{g \times 2r}[s, s^{-1}]$. Then there exists a quaternionic matrix $\widetilde{R}$ such that $\mathcal{B} = \ker \widetilde{R}\,(\sigma, \sigma^{-1})$.*

Proof. Let $\mathcal{B}^{\mathbb{C}} = \ker R$. Partition the matrix R as $R = \begin{bmatrix} R_1 & R_2 \end{bmatrix}$, where $R_1, R_2 \in \mathbb{C}^{g \times r}[s, s^{-1}]$, and construct the quaternionic matrix $\widetilde{R} = R_1 + R_2 \mathbf{j}$. We claim that $\mathcal{B} = \ker \widetilde{R}$, or equivalently, by Proposition 4.2.6, that

$$\mathcal{B}^{\mathbb{C}} = \ker \widetilde{R}^c, \quad \text{with} \quad \widetilde{R}^c = \begin{bmatrix} R_1 & R_2 \\ -\overline{R}_2 & \overline{R}_1 \end{bmatrix}.$$

In order to prove this we just need to see that $\ker R = \ker \widetilde{R}^c$. It is obvious that $\ker \widetilde{R}^c \subseteq \ker R$. Consider now a trajectory $v \in \ker R$. By definition of $\mathcal{B}^{\mathbb{C}}$ there exists $w = w_1 + w_2 \mathbf{j} \in \mathcal{B}$ such that $v = w^{\mathbb{C}} = \left[\begin{smallmatrix} w_1 \\ -\overline{w}_2 \end{smallmatrix}\right]$. To show that $v \in \ker \widetilde{R}^c$ it suffices to prove that $\begin{bmatrix} -\overline{R}_2 & \overline{R}_1 \end{bmatrix} v = 0$, i.e., $\overline{R}_2 w_1 + \overline{R}_1 \overline{w}_2 = 0$. Since $\mathcal{B}$ is linear on the right, then $\widetilde{w} = -w\mathbf{j} = w_2 - w_1 \mathbf{j} \in \mathcal{B}$. Therefore $\widetilde{w}^{\mathbb{C}} = \left[\begin{smallmatrix} w_2 \\ \overline{w}_1 \end{smallmatrix}\right] \in \mathcal{B}^{\mathbb{C}}$, i.e., $R_1 w_2 + R_2 \overline{w}_1 = 0$. Hence, $\overline{R}_1 \overline{w}_2 + \overline{R}_2 w_1 = 0$ by conjugation. □

Propositions 4.2.6 and 4.2.7 show that the analysis of $\mathcal{B}$ is equivalent to the analysis of its complex form $\mathcal{B}^{\mathbb{C}}$. In the same way, it was shown in Section 3.3 that quaternionic polynomial matrices share many algebraic properties with their complex adjoint matrices. These equivalences will play an important role in the sequel, where we investigate a fundamental equivalence relation for kernel representations.

Definition 4.2.8. Let $R_l \in \mathbb{H}^{g_l \times r}[s, s^{-1}]$, $l = 1, 2$. Then R_1 and R_2 are said to be *equivalent representations* if $\ker R_1\,(\sigma, \sigma^{-1}) = \ker R_2\,(\sigma, \sigma^{-1})$.

Example 4.2.9. Consider the following quaternionic polynomial matrices

$$R_1 = \begin{bmatrix} s & -\mathbf{i} \\ 0 & s-\mathbf{k} \end{bmatrix}, \quad R_2 = \begin{bmatrix} s+\mathbf{k} & 0 \\ \mathbf{j} & 1 \end{bmatrix}. \tag{4.5}$$

These are equivalent representations of the same behavior which, as it is easy to check, is

$$\ker R_1 = \ker R_2 = \left\{ \begin{bmatrix} \mathbf{j} \\ 1 \end{bmatrix} \mathbf{k}^t q,\ q \in \mathbb{H} \right\}.$$

A straightforward calculation shows that $R_1 = UR_2$ and $R_2 = U^{-1}R_1$, where

$$U = \begin{bmatrix} 1 & -\mathbf{i} \\ -\mathbf{j} & s-\mathbf{k} \end{bmatrix}$$

is an unimodular L-polynomial matrix. □

We will show that, as in the real and in the complex cases, two representations are equivalent if and only if each one is a left multiple of the other, as in the previous example. This main result is a consequence of the following more general statement.

Theorem 4.2.10. *Let R_1 and R_2 be two quaternionic L-polynomial matrices. Then* $\ker R_1 \subseteq \ker R_2$ *if and only if there exists a quaternionic L-polynomial matrix X such that $XR_1 = R_2$.*

Proof. "If" part. Assume that $R_2 = XR_1$ and let $w \in \ker R_1$. Then, $R_2 w = XR_1 w = 0$ and therefore $w \in \ker R_2$ showing that $\ker R_1 \subseteq \ker R_2$.

"Only if" part. Assume now that $\ker R_1 \subseteq \ker R_2$. We wish to prove that there exists a matrix X such that $XR_1 = R_2$. By Proposition 4.2.6,

$$\ker R_1 \subseteq \ker R_2 \Leftrightarrow \ker R_1^c \subseteq \ker R_2^c.$$

Similarly to what is stated in [52, Section 4] for the real case, this implies that there exists a complex polynomial matrix Y such that

$$YR_1^c = R_2^c.$$

By Proposition 3.3.3, this means that there also exists a quaternionic matrix X such that

$$XR_1 = R_2,$$

thus proving the theorem. $\square$

Corollary 4.2.11. *Two quaternionic representations $R_1(s, s^{-1})$, $R_2(s, s^{-1})$ are equivalent if and only if there exist $X_1(s, s^{-1})$ and $X_2(s, s^{-1})$ such that $R_1 = X_1 R_2$ and $R_2 = X_2 R_1$. Moreover, if both matrices are frr then $X_1 = X_2^{-1}$, i.e., X_1 and X_2 are unimodular matrices.*

Proof. The first part of the corollary is a trivial consequence of Theorem 4.2.10.

Suppose now that R_1 and R_2 are frr. Since $R_1 = X_1R_2$ and $R_2 = X_2R_1$, then $R_1 = X_1X_2R_1$ and $R_2 = X_2X_1R_2$. Hence we have that $X_1X_2 = X_2X_1 = I$, i.e., $X_1 = X_2^{-1}$ and X_1 and X_2 are unimodular. $\square$

Remark 4.2.12. Since s^l is an invertible element in $\mathbb{H}[s, s^{-1}]$, it follows that

$$\ker R(\sigma, \sigma^{-1}) = \ker \sigma^l R(\sigma, \sigma^{-1}).$$

As a consequence, it is always possible to choose a polynomial representation for any behavior $\mathcal{B}$. Indeed, if $\mathcal{B}$ has a representation $R(s, s^{-1})$, then, for an adequate integer $M \geq 0$, $s^M R(s, s^{-1}) \in \mathbb{H}^{g\times r}[s]$ is still a representation of $\mathcal{B}$. Therefore, for the sake of simplicity, we shall choose polynomial kernel representations, although always regarding them as L-polynomial matrices. $\square$

4.2.2 Hybrid representations

Although the representation of a behavior as a kernel is very general, it is sometimes possible or desirable to use other representations which, besides the system variables, involve also latent variables, often introduced from first principle modelling or in order to simplify the description of a system. After formally defining such representations we prove that kernel behaviors are the ones and only ones that admit a hybrid representation.

Definition 4.2.13. Consider a system $\Sigma = (\mathbb{Z}, \mathbb{H}^r, \mathcal{B})$. The equation

$$R\left(\sigma, \sigma^{-1}\right) w = M\left(\sigma, \sigma^{-1}\right) a, \tag{4.6}$$

where $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ and $M \in \mathbb{H}^{g\times l}[s, s^{-1}]$, is said to be a *hybrid representation* of $\mathcal{B}$ if

$$\mathcal{B} = \left\{w \in (\mathbb{H}^r)^{\mathbb{Z}} \,:\, \exists\, a \in (\mathbb{H}^l)^{\mathbb{Z}} \text{ such that } w \text{ and } a \text{ satisfy (4.6)}\right\}.$$

The variables a are called latent variables. The set

$$\mathcal{B}^{(w,a)} = \left\{(w, a) \in (\mathbb{H}^{r+l})^{\mathbb{Z}} \,:\, R\left(\sigma, \sigma^{-1}\right) w = M\left(\sigma, \sigma^{-1}\right) a\right\}$$

is said to be the *full behavior* of Σ and $\mathcal{B}$ is said to be the *external behavior* of $\mathcal{B}^{(w,a)}$.

In particular, when the matrix R is the identity matrix, then the behavior $\mathcal{B}$ of (4.6) is given as the image of the quaternionic polynomial matrix operator $M(\sigma, \sigma^{-1})$, i.e., $\mathcal{B} = \operatorname{im} M(\sigma, \sigma^{-1})$. Representations of this form are called *image representations*. Behaviors which admit an image representation will be characterized in terms of controllability properties in Section 5.1.

Note that every behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ which admits a kernel representation, i.e., $\mathcal{B} = \ker R$ for some $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$, admits also a hybrid representation of the form (4.6). Indeed, it is enough to consider $M(s, s^{-1}) = 0$. In the sequel it is shown that, analogously to the commutative case [43], the converse also holds, i.e., a behavior which admits a hybrid representation admits as well a kernel representation. First we show that any nonzero quaternionic polynomial operator is surjective and that a quaternionic polynomial matrix operator $R(\sigma, \sigma^{-1})$ is surjective if and only if the matrix $R(s, s^{-1})$ has full row rank.

Lemma 4.2.14. *Let $p(s)$ be a nonzero quaternionic polynomial of degree n. Then the operator $p(\sigma)$ is surjective.*

Proof. As mentioned in section 2.2.2, each polynomial can be decomposed in linear factors and therefore there exist quaternions $\alpha_1, \ldots, \alpha_n$ such that $p(s) = (s-\alpha_n)\cdots(s-\alpha_1)$. Hence it is enough to prove that the polynomial $q(s) = s - \alpha$, $\alpha \in \mathbb{H}$, is surjective, i.e., $\forall\, w \in (\mathbb{H})^{\mathbb{Z}} :\ \exists\, v \in (\mathbb{H})^{\mathbb{Z}}$ such that $w = qv$. Given any $w(\cdot) \in (\mathbb{H})^{\mathbb{Z}}$ and $\beta \in \mathbb{H}$ construct $v(\cdot) \in (\mathbb{H})^{\mathbb{Z}}$ by the recursive formula

$$v(0) = \beta, \qquad v(t) = \alpha^t v(0) + \sum_{m=0}^{t-1} \alpha^{t-1-m} w(m).$$

It is not difficult to check that w and v satisfy $w = qv$ and the result follows. □

Lemma 4.2.15. *A matrix $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ has full row rank if and only if the operator $R\,(\sigma, \sigma^{-1})$ is surjective.*

Proof. "If" part. Suppose that the matrix R has not frr. Then, by Theorem 3.1.11 there exists an unimodular matrix $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ such that $UR = \left[\begin{smallmatrix} R' \\ 0 \end{smallmatrix}\right]$, with R' frr. Partition conformably the matrix U as $U = \left[\begin{smallmatrix} U' \\ U'' \end{smallmatrix}\right]$. Thus

$$w = Rv \Leftrightarrow Uw = URv \Leftrightarrow \begin{bmatrix} U' \\ U'' \end{bmatrix} w = \begin{bmatrix} R' \\ 0 \end{bmatrix} v,$$

which implies that $U''w = 0$, i.e., the equation $w = Rv$ only has solution if $w \in \ker U'' \neq (\mathbb{H}^g)^{\mathbb{Z}}$ and therefore $R\,(\sigma, \sigma^{-1})$ is not surjective.

"Only if" part. Since the matrix R has frr, then by Theorem 3.4.3 there exist unimodular matrices $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ and $V \in \mathbb{H}^{r\times r}[s, s^{-1}]$ such that

$$URV = \Gamma = \left[\begin{array}{ccc|c} \gamma_1 & & & \\ & \ddots & & 0 \\ & & \gamma_g & \end{array}\right],$$

where the polynomials $\gamma_l \in \mathbb{H}[s]$, $l = 1, \ldots, g$, are monic. Let $\widetilde{w} = Uw = \begin{bmatrix} \widetilde{w}_1 & \cdots & \widetilde{w}_g \end{bmatrix}^T$ and $\widetilde{v} = V^{-1}v = \begin{bmatrix} \widetilde{v}_1 & \cdots & \widetilde{v}_g \end{bmatrix}^T$. Then

$$w = Rv \Leftrightarrow Uw = URVV^{-1}v \Leftrightarrow \widetilde{w} = \Gamma\widetilde{v} \Leftrightarrow \begin{cases} \widetilde{w}_1 &= \gamma_1\widetilde{v}_1 \\ & \vdots \\ \widetilde{w}_g &= \gamma_g\widetilde{v}_g \end{cases},$$

and therefore the operator $R\,(\sigma, \sigma^{-1})$ is surjective if and only if each polynomial operator $\gamma\,(\sigma)$ is surjective, which is indeed the case by Lemma 4.2.14. □

We are now in a position to show that a behavior which admits a representation with latent variables admits as well a kernel representation. The procedure presented in the following proposition is usually called *elimination of latent variables* [43].

Proposition 4.2.16. *Let $R \in \mathbb{H}^{g\times r}[s,s^{-1}]$, $M \in \mathbb{H}^{g\times l}[s,s^{-1}]$ and assume that the behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ admits the representation*

$$R\left(\sigma,\sigma^{-1}\right) w = M\left(\sigma,\sigma^{-1}\right) a.$$

Let $U \in \mathbb{H}^{g\times g}[s,s^{-1}]$ be a unimodular matrix such that

$$UM = \begin{bmatrix} M' \\ 0 \end{bmatrix}, \quad UR = \begin{bmatrix} R' \\ R'' \end{bmatrix},$$

with $M'(s,s^{-1})$ a full row rank matrix. Then

$$\mathcal{B} = \ker R''\left(\sigma,\sigma^{-1}\right).$$

Remark 4.2.17. Note that such a unimodular matrix U always exists by Theorem 3.1.11. □

Proof. By hypothesis we have that

$$\begin{aligned} \mathcal{B} &= \left\{w \in (\mathbb{H}^r)^{\mathbb{Z}} \,:\, \exists\, a \in (\mathbb{H}^l)^{\mathbb{Z}} \text{ such that } Rw = Ma\right\} \\ &= \left\{w \in (\mathbb{H}^r)^{\mathbb{Z}} \,:\, \exists\, a \in (\mathbb{H}^l)^{\mathbb{Z}} \text{ such that } URw = UMa\right\} \\ &= \left\{w \in (\mathbb{H}^r)^{\mathbb{Z}} \,:\, \exists\, a \in (\mathbb{H}^l)^{\mathbb{Z}} \text{ such that } R'w = M'a \text{ and } R''w = 0\right\}. \end{aligned}$$

Since the matrix M' has frr, by Lemma 4.2.15 the operator $M'\left(\sigma,\sigma^{-1}\right)$ is surjective. Then in particular for all $w \in \ker R''$ there exists an $a \in (\mathbb{H}^l)^{\mathbb{Z}}$ such that $R'w = M'a$. Hence, $\mathcal{B} = \ker R''\left(\sigma,\sigma^{-1}\right)$. □

In the following example we illustrate the elimination of variables in the quaternionic case.

Example 4.2.18. Let

$$R = \begin{bmatrix} s & \mathbf{i} \\ -1 & -3\mathbf{i}s^{-1} \\ -\mathbf{j}s & s \end{bmatrix} \in \mathbb{H}^{3\times 2}[s,s^{-1}] \quad , \quad M = \begin{bmatrix} 1 \\ -2s^{-1} \\ -\mathbf{j} \end{bmatrix} \in \mathbb{H}^{3\times 1}[s,s^{-1}]$$

and assume that

$$R\left(\sigma,\sigma^{-1}\right)w = M\left(\sigma,\sigma^{-1}\right)a$$

is a hybrid representation of the behavior $\mathcal{B} \subseteq (\mathbb{H}^2)^{\mathbb{Z}}$. The unimodular L-polynomial matrix

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 2 & s & 0 \\ \mathbf{j} & 0 & 1 \end{bmatrix} \in \mathbb{H}^{3\times 3}[s, s^{-1}]$$

is such that

$$UR = \begin{bmatrix} s & \mathbf{i} \\ s & -\mathbf{i} \\ 0 & s-\mathbf{k} \end{bmatrix} \quad \text{and} \quad UM = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Therefore, by Proposition 4.2.16, $\mathcal{B} = \ker \begin{bmatrix} s & -\mathbf{i} \\ 0 & s-\mathbf{k} \end{bmatrix}$ and, by Example 4.2.9,

$$\mathcal{B} = \left\{ \begin{bmatrix} \mathbf{j} \\ 1 \end{bmatrix} \mathbf{k}^t q, \ q \in \mathbb{H} \right\}.$$

□

4.2.3 Input-output representations

In section 4.2.1 we considered behaviors that were the kernel of a difference operator $R\left(\sigma,\sigma^{-1}\right)$, for some $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$. When the rank of the matrix R is less than the number of columns it is possible to write an input/output (i/o) representation of the behavior.

To introduce the class of i/o systems in a proper way, we need the following preliminary definition.

Definition 4.2.19. Let $\mathcal{B} \subseteq (\mathbb{H}^q)^{\mathbb{Z}}$, consider a variable u consisting of m components of the system variable w and a variable y consisting of the remaining $p = q - m$ components. Then u is an *input variable* and y is an *output variable* of $\mathcal{B}$ if

1. u is *free* in $\mathcal{B}$, i.e., $\forall u \in (\mathbb{H}^m)^{\mathbb{Z}}$, $\exists y \in (\mathbb{H}^p)^{\mathbb{Z}}$ such that $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathcal{B}$;

2. once u is fixed, no component of y is free in $\{y : \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathcal{B}\}$.

Let $\mathcal{B} = \ker R$. Partitioning the components of the trajectories $w \in \mathcal{B}$ into input and output variables, i.e., considering a permutation matrix Π such that $\Pi w = \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$, where u is an input variable and y is an output variable, will determine a partition of the columns of R

$$R\Pi^{-1} = \begin{bmatrix} P & -Q \end{bmatrix}.$$

Thus

$$Rw = 0 \;\Leftrightarrow\; R\Pi^{-1}\Pi w = 0 \;\;\Leftrightarrow\; \begin{cases} \begin{bmatrix} P & -Q \end{bmatrix} \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] = 0 \\ w = \Pi^{-1} \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \end{cases}$$

$$\Leftrightarrow \begin{cases} Py = Qu \\ w = \Pi^{-1} \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \end{cases}, \tag{4.7}$$

(4.7) is called the *i/o representation* of the behavior $\mathcal{B}$.

An *i/o behavior* is a behavior whose variables are partitioned into inputs and outputs and is denoted by $\mathcal{B}_{\mathrm{i/o}}$. Here we consider i/o behaviors given by $\mathcal{B}_{\mathrm{i/o}} = \ker R$ with $R = \begin{bmatrix} P & -Q \end{bmatrix}$, i.e.,

$$\mathcal{B}_{\mathrm{i/o}} = \left\{ \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \left(\mathbb{H}^{p+m}\right)^{\mathbb{Z}} \;:\; Py = Qu \right\}, \tag{4.8}$$

where u is an input variable and y is an output variable. We will say that (P, Q) is an i/o representation of $\mathcal{B}_{\mathrm{i/o}}$. Note that the complex form of the behavior $\mathcal{B}_{\mathrm{i/o}}$, $\mathcal{B}^{\mathbb{C}}_{\mathrm{i/o}}$, is also an i/o behavior which, after a suitable permutation of its variables, has kernel representation $\begin{bmatrix} P^c & -Q^c \end{bmatrix}$.

Note that the fact that u is an input (and y is an output) is equivalent to say that P has full column rank (fcr) and $\operatorname{im} Q \subset \operatorname{im} P$.

If P has not frr, pre-multiplication by a suitable unimodular matrix U yields

$$UP = \left[\begin{smallmatrix} \tilde{P} \\ 0 \end{smallmatrix}\right]$$

with $\tilde{P}$ square and full rank. But, since $\operatorname{im} Q \subset \operatorname{im} P$, also

$$UQ = \left[\begin{smallmatrix} \tilde{Q} \\ 0 \end{smallmatrix}\right]$$

where $\tilde{Q}$ has the same number of rows as $\tilde{P}$.

This means that the i/o behavior with i/o representation (P, Q) has an equivalent representation $\left(\tilde{P}, \tilde{Q}\right)$ with $\tilde{P}$ square and full rank, i.e., with $\tilde{P}$ invertible (over the quaternionic rational matrices). For this reason, without loss of generality, we will henceforth only consider i/o representations (P, Q) with invertible P.

Moreover, we will only deal with proper systems, which means that we also assume that the *transfer matrix* $P^{-1}Q$ of the behavior (4.8) is a proper rational matrix, i.e., in each entry the degree of the numerator does not exceed the degree of the denominator.

Definition 4.2.20. A dynamical system defined by the equation

$$Py = Qu, \tag{4.9}$$

where $P \in \mathbb{H}^{p\times p}[s, s^{-1}]$ and $Q \in \mathbb{H}^{p\times m}[s, s^{-1}]$, is a *(proper) quaternionic i/o system*, with behavior $\mathcal{B}_{\mathrm{i/o}}$ defined by equation (4.8), if P admits a rational inverse and its transfer matrix $P^{-1}Q \in \mathbb{H}^{p\times m}(s, s^{-1})$ is proper.

4.3 The solution of quaternionic dynamical equations

The main goal of this section is to give a complete and explicit characterization of all solutions of the quaternionic matrix dynamical equations. We consider first the discrete case, i.e., we characterize the solutions of the quaternionic matrix difference equation

$$R_p w(t+p) + \cdots + R_1 w(t+1) + R_0 w(t) = 0, \tag{4.10}$$

where $R_l \in \mathbb{H}^{g\times r}, \quad l = 0, \ldots, p.$

Using the previously defined shift operator, (4.10) can be written in the form

$$R(\sigma)w = 0, \tag{4.11}$$

with $R(s) = R_p s^p + \cdots + R_1 s + R_0$.

4.3.1 The First Order Case

We will first consider the case where $R(s) = I_g s - A,\ A \in \mathbb{H}^{g\times g}$, i.e., we will give the solutions of the first order equation

$$w(t+1) = Aw(t). \tag{4.12}$$

As happens in the commutative case, the solutions of (4.12) are clearly $w(t) = A^t w(0)$. Since, in general, it is not easy to compute the powers of the matrix A, we present next a method to calculate explicitly the solutions of (4.12) without computing the powers of A.

Since $\big(w(t)\big)^{\mathbb{C}} = w^{\mathbb{C}}(t)$, applying Proposition 1.2.3-4 we have that w is a solution of (4.12) if and only if $w^{\mathbb{C}}$ is a solution of

$$w^{\mathbb{C}}(t+1) = A^c w^{\mathbb{C}}(t). \tag{4.13}$$

We will compute the solutions of (4.13) in the following way.

First we calculate $\sigma(A^c) = \{\lambda_1, \ldots, \lambda_g, \overline{\lambda}_1, \ldots, \overline{\lambda}_g\}$.

By Lemma 1.2.25, the (generalized) eigenvector matrix W is a complex adjoint matrix, i.e., $W = V^c$ for some $V \in \mathbb{H}^{g\times g}$. Therefore, we just need to calculate the eigenvectors (or generalized eigenvectors) associated with the first g eigenvalues of $\sigma(A^c)$. In fact, if these vectors form a matrix $\left[\begin{smallmatrix} V_1 \\ -\overline{V}_2 \end{smallmatrix}\right]$, then W will be of the form $W = \left[\begin{smallmatrix} V_1 & V_2 \\ -\overline{V}_2 & \overline{V}_1 \end{smallmatrix}\right]$.

The Jordan form of A^c is given by $(V^c)^{-1}A^cV^c$. It obviously is a complex adjoint matrix and, being an upper triangular matrix, it has the form

$$J^c = \left[\begin{array}{c|c} J & 0 \\ \hline 0 & \overline{J} \end{array}\right].$$

Then

$$w^{\mathbb{C}}(t) = (A^c)^t w^{\mathbb{C}}(0) = V^c (J^c)^t (V^c)^{-1} w^{\mathbb{C}}(0) = V^c (J^c)^t q^{\mathbb{C}},$$

with $q^{\mathbb{C}} = \begin{bmatrix} q_1 \\ -\overline{q}_2 \end{bmatrix} = (V^c)^{-1} w^{\mathbb{C}}(0) \in \mathbb{C}^{2g}$.

The solutions w of (4.12) are of the form

$$\boxed{w(t) = VJ^t q}, \qquad \begin{array}{c} q = q_1 + q_2\mathbf{j} \\ V = V_1 + V_2\mathbf{j} \end{array}, \quad q \in \mathbb{H}^g.$$

Example 4.3.1. In this example we will calculate the solutions of the equation $w(t+1) = Aw(t)$ where

$$A = \begin{bmatrix} 1+2\mathbf{i} & \mathbf{j} \\ -\mathbf{j} & 1-2\mathbf{i} \end{bmatrix}.$$

The complex adjoint matrix of A is

$$A^c = \begin{bmatrix} 1+2\mathbf{i} & 0 & 0 & 1 \\ 0 & 1-2\mathbf{i} & -1 & 0 \\ 0 & -1 & 1-2\mathbf{i} & 0 \\ 1 & 0 & 0 & 1+2\mathbf{i} \end{bmatrix}$$

and

$$\sigma(A^c) = \{\pm 2\mathbf{i}, 2 \pm 2\mathbf{i}\};$$

the eigenvector matrix is

$$V^c = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Therefore

$$w^{\mathbb{C}}(t) = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} (2\mathbf{i})^t & 0 & 0 & 0 \\ 0 & (2+2\mathbf{i})^t & 0 & 0 \\ 0 & 0 & (-2\mathbf{i})^t & 0 \\ 0 & 0 & 0 & (2-2\mathbf{i})^t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ -\overline{c}_3 \\ -\overline{c}_4 \end{bmatrix}$$

and

$$w(t) = \begin{bmatrix} -1 & 1 \\ -\mathbf{j} & -\mathbf{j} \end{bmatrix} \begin{bmatrix} (2\mathbf{i}^t) & 0 \\ 0 & (2+2\mathbf{i})^t \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix},$$

where $\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} c_1 + c_3\mathbf{j} \\ c_2 + c_4\mathbf{j} \end{bmatrix}$, $c_1,\ c_2,\ c_3,\ c_4 \in \mathbb{C}$. □

Remark 4.3.2. The solution of the first order differential equation

$$\dot{w}(t) = Aw(t), \quad A \in \mathbb{H}^{g\times g} \tag{4.14}$$

can be calculated by a similar procedure. Indeed, again by Proposition 1.2.3-4 we have that w is a solution of (4.14) if and only if $w^{\mathbb{C}}$ is a solution of

$$\dot{w}^{\mathbb{C}}(t) = A^c w^{\mathbb{C}}(t). \tag{4.15}$$

It is well-know (see for instance [43]) that the solutions of equation (4.15) are of the form

$$w^{\mathbb{C}}(t) = e^{A^c t} w^{\mathbb{C}}(0), \qquad e^{A^c t} = \sum_{r=0}^{\infty} \frac{(A^c)^r\, t^r}{r!}. \tag{4.16}$$

If, with the notation given above, the Jordan form of A^c is $J^c = (V^c)^{-1}A^cV^c$, equation (4.16) can be written as

$$w^{\mathbb{C}}(t) = V^c e^{J^c t} q^{\mathbb{C}}$$

and therefore the solutions w of (4.14) are of the form

$$w(t) = Ve^{Jt}q.$$

□

For this reason, in the cases to be considered in the sequel we will only give the solutions of quaternionic difference equations.

4.3.2 The Higher Order Case

In this section we will give the solutions of the matrix difference equation (4.10) in the higher order case, i.e., $p > 1$. First we will assume that $R_p = I_g$ and consider the equation

$$w(t+p) + R_{p-1}w(t+p-1) + \cdots + R_0 w(t) = 0, \tag{4.17}$$

where $R_l \in \mathbb{H}^{g\times g}, \quad l = 0, \ldots, p-1$.

As usual, by introducing latent variables this equation can be reduced to the first order case in the following way:

$$\begin{cases} x_1(t) &= w(t) \\ x_2(t) &= w(t+1) \\ \vdots & \vdots \\ x_{p-1}(t) &= w(t+p-2) \\ x_p(t) &= w(t+p-1) \end{cases}$$

which is equivalent to

$$\underbrace{\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ \vdots \\ x_{p-1}(t+1) \\ x_p(t+1) \end{bmatrix}}_{x(t+1)} = \underbrace{\begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ -R_0 & -R_1 & -R_2 & \cdots & -R_{p-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{p-1}(t) \\ x_p(t) \end{bmatrix}}_{x(t)},$$

or, in matrix notation,

$$x(t+1) = Ax(t)\ ,\ A \in \mathbb{H}^{pg\times pg}. \tag{4.18}$$

In order to obtain an explicitly form for the solutions of (4.18), the procedure of the previous section can be applied. Since

$$w(t) = \begin{bmatrix} I_g & 0 & \cdots & 0 \end{bmatrix} x(t),$$

this yields an explicitly form for the solutions of the initial equation.

Example 4.3.3. In this example the solutions of the (scalar) second order difference equation

$$w(t+2) - (\mathbf{j}+\mathbf{i})w(t+1) - \mathbf{k}w(t) = 0 \tag{4.19}$$

are obtained. Reducing it to a first order matrix difference equation

$$x(t+1) = \begin{bmatrix} 0 & 1 \\ \mathbf{k} & \mathbf{j}+\mathbf{i} \end{bmatrix} x(t),$$

where $w(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t)$ and applying the procedure of section 4.3.1 we have

$$x(t) = \begin{bmatrix} -2\mathbf{i} \\ 2 \end{bmatrix} \mathbf{i}^t q_1 + \left(\begin{bmatrix} 3-\mathbf{k} \\ \mathbf{i}-\mathbf{j} \end{bmatrix} \mathbf{i}^t + \begin{bmatrix} -2\mathbf{i} \\ 2 \end{bmatrix} t\mathbf{i}^{t-1} \right) q_2, \quad q_1, q_2 \in \mathbb{H}$$

and therefore the general solution of (4.19) is given by

$$w(t) = -2\mathbf{i}^{t+1} q_1 + (3 - \mathbf{k} - 2t)\mathbf{i}^t q_2, \quad q_1, q_2 \in \mathbb{H}.$$

$\square$

Assume now that the matrix R_p in equation (4.10) is a non invertible square matrix and that $R(s) = R_p s^p + \cdots + R_1 s + R_0$ has frr.

By Theorem 3.1.12, there exists a unimodular matrix $U(s) \in \mathbb{H}^{g\times g}[s]$ such that

$$U(s)R(s) = \operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) \widetilde{R}_h + \widetilde{R}(s),$$

where $\widetilde{R}_h \in \mathbb{H}^{g\times g}$ has frr (and hence invertible) and each row degree of $\widetilde{R}(s)$ is strictly smaller than the corresponding row degree of $\operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) \widetilde{R}_h$. Let

$$h = \max m_l, \; d_l = h - m_l \;\text{ and }\; W(s) = \operatorname{diag}\left(s^{d_1}, \ldots, s^{d_g}\right), \; l = 1, \ldots, g.$$

Then

$$\begin{aligned} \widetilde{R}_h^{-1} W(s)U(s)R(s) &= \widetilde{R}_h^{-1} \operatorname{diag}\left(s^{d_1}, \ldots, s^{d_g}\right) \left[\operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) \widetilde{R}_h + \widetilde{R}(s)\right] \\ &= I_g s^h + \widetilde{R}_h^{-1} W(s) \widetilde{R}(s) \\ &= \underbrace{I_g s^h + \widetilde{R}_h^{-1}\left(\widetilde{R}_{h-1} s^{h-1} + \cdots \widetilde{R}_1 s + \widetilde{R}_0\right)}_{R'(s)} \end{aligned}$$

Since $W(s)$ is unimodular as an element of $\mathbb{H}^{g\times g}[s, s^{-1}]$ we have

$$R(\sigma)w = 0 \Leftrightarrow R'(\sigma)w = 0.$$

Since the equation on the right-hand side is of the same form as (4.17), its solutions are obtained by a similar procedure.

4.3.3 The Non Square Case

Finally, consider equation (4.10), now with $g < r$ and $R(s) = R_p s^p + \cdots + R_1 s + R_0$ frr.

By Theorem 3.1.12, there exists a unimodular matrix $U(s) \in \mathbb{H}^{g\times g}[s]$ such that

$$U(s)R(s) = \operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) \widetilde{R}_n + \operatorname{diag}\left(s^{m_1-1}, \ldots, s^{m_g-1}\right) \widetilde{R}_{n-1} + \cdots, \tag{4.20}$$

where $\widetilde{R}_n$ has frr. This implies that there exists a permutation matrix $\Pi \in \mathbb{H}^{r\times r}$ such that

$$\widetilde{R}_n \Pi = [\widetilde{P}_n \mid -\widetilde{Q}_n],$$

where $\widetilde{P}_n \in \mathbb{H}^{g\times g}$ is invertible. Without loss of generality, we will assume that $\widetilde{P}_n = I_g$. Partitioning $\Pi^{-1}w$ conformably as

$$\Pi^{-1}w = \begin{bmatrix} y \\ u \end{bmatrix},$$

yields

$$\begin{gathered} R(\sigma)w = 0 \\ \Updownarrow \\ \left(\operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) + \operatorname{diag}\left(s^{m_1-1}, \ldots, s^{m_g-1}\right) \widetilde{P}_{n-1} + \cdots\right) y \\ = \left(\operatorname{diag}\left(s^{m_1}, \ldots, s^{m_g}\right) \widetilde{Q}_n + \cdots\right) u \\ \Updownarrow \end{gathered}$$

$$\left(\sigma^n + \sigma^{n-1}\widetilde{P}_{n-1} + \cdots + \sigma\widetilde{P}_1 + \widetilde{P}_0\right) y = \left(\sigma^n \widetilde{Q}_n + \cdots + \sigma\widetilde{Q}_1 + \widetilde{Q}_0\right) u, \tag{4.21}$$

where $n = \max m_l,\ l = 1, \ldots, g$.

Defining the vector x with components

$$\begin{cases} x_n & = & y - \widetilde{Q}_n u \\ x_{n-1} & = & \sigma x_n + \widetilde{P}_{n-1} y - \widetilde{Q}_{n-1} u \\ x_{n-2} & = & \sigma x_{n-1} + \widetilde{P}_{n-2} y - \widetilde{Q}_{n-2} u \\ & \vdots & \\ x_1 & = & \sigma x_2 + \widetilde{P}_1 y - \widetilde{Q}_1 u \end{cases}, \tag{4.22}$$

we obtain from (4.21) and (4.22)

$$\begin{cases} \sigma x_1 & = & -\widetilde{P}_0 x_n + (\widetilde{Q}_0 - \widetilde{P}_0 \widetilde{Q}_n) u \\ \sigma x_2 & = & x_1 - \widetilde{P}_1 x_n + (\widetilde{Q}_1 - \widetilde{P}_1 \widetilde{Q}_n) u \\ \sigma x_3 & = & x_2 - \widetilde{P}_2 x_n + (\widetilde{Q}_2 - \widetilde{P}_2 \widetilde{Q}_n) u \\ & \vdots & \\ \sigma x_n & = & x_{n-1} - \widetilde{P}_{n-1} x_n + (\widetilde{Q}_{n-1} - \widetilde{P}_{n-1} \widetilde{Q}_n) u \\ y & = & x_n + \widetilde{Q}_n u \end{cases}, \tag{4.23}$$

which is equivalent to

$$\begin{cases} \sigma x & = & \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 & -\widetilde{P}_0 \\ I & 0 & \cdots & \cdots & \cdots & -\widetilde{P}_1 \\ 0 & I & 0 & \cdots & \cdots & -\widetilde{P}_2 \\ \vdots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & I & 0 & -\widetilde{P}_{n-2} \\ 0 & \cdots & \cdots & 0 & I & -\widetilde{P}_{n-1} \end{bmatrix} x + \begin{bmatrix} \widetilde{Q}_0 - \widetilde{P}_0 \widetilde{Q}_n \\ \widetilde{Q}_1 - \widetilde{P}_1 \widetilde{Q}_n \\ \vdots \\ \vdots \\ \widetilde{Q}_{n-2} - \widetilde{P}_{n-2} \widetilde{Q}_n \\ \widetilde{Q}_{n-1} - \widetilde{P}_{n-1} \widetilde{Q}_n \end{bmatrix} u \\ y & = & \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 & I \end{bmatrix} x + \widetilde{Q}_n u \end{cases}, \tag{4.24}$$

and can be written as

$$\begin{cases} \sigma x & = & Ax + Bu \\ y & = & Cx + Du \end{cases}, \tag{4.25}$$

with A, B, C, D defined in the obvious way.

The trajectories w are given by $\Pi \left[\begin{smallmatrix} y \\ u \end{smallmatrix} \right]$, where $\left[\begin{smallmatrix} y \\ u \end{smallmatrix} \right]$ are the trajectories obtained from (4.25).

Thus our problem amounts to find the solutions of the first order non homogeneous equation

$$x(t+1) = Ax(t) + Bu(t). \tag{4.26}$$

It follows from Proposition 1.2.3-4 that x is a solution of (4.26) if and only if $x^{\mathbb{C}}$ is a solution of

$$x^{\mathbb{C}}(t+1) = A^c x^{\mathbb{C}}(t) + B^c u^{\mathbb{C}}(t). \tag{4.27}$$

Since, as is well known, the solutions of (4.27) are given by

$$x^{\mathbb{C}}(t) = (A^c)^t \, x^{\mathbb{C}}(0) + \sum_{m=0}^{t-1} (A^c)^{t-1-m} \, B^c u^{\mathbb{C}}(m), \tag{4.28}$$

the solutions of (4.26) are of the form

$$x(t) = A^t x(0) + \sum_{m=0}^{t-1} A^{t-1-m} B u(m). \tag{4.29}$$

The computation of the powers A^t can be made via the Jordan form, as shown in Section 4.3.1.

Chapter 5

Dynamical properties of quaternionic behaviors

In this chapter we show how basic but fundamental dynamical properties of a quaternionic behavior can be characterized in terms of its (kernel) representations. It turns out that the algebraic tools introduced before, such as the quaternionic Smith and Smith-McMillan forms and the polynomial determinant Pdet, play an important role in these characterizations.

5.1 Controllability

The concept of controllability plays a fundamental role in systems theory. Roughly speaking, we call a behavior controllable if it is possible to switch freely from one to another of its trajectories in finite time. We show that many usual characterizations of controllability also hold in the quaternionic case, while the well-known Hautus criterion does not. We start by giving the definition of controllable behaviors.

Definition 5.1.1. [43] A behavior $\mathcal{B}$ of a time-invariant dynamical system is called *controllable* if for any two trajectories $w_1, w_2 \in \mathcal{B}$, and any time instant t_1, there exists

$t_2 > t_1$ and a trajectory $w \in \mathcal{B}$ such that

$$w(t) = \begin{cases} w_1(t), & t \leq t_1; \\ w_2(t), & t \geq t_2. \end{cases} \tag{5.1}$$

When property (5.1) holds, w_1 and w_2 are said to be *concatenable* in $\mathcal{B}$. Therefore $\mathcal{B}$ is controllable if all its trajectories are concatenable in $\mathcal{B}$.

In the sequel we shall characterize the controllability of discrete-time systems.

Lemma 5.1.2. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ and $\mathcal{B} = \ker R$. Then $\mathcal{B}$ is controllable if and only if $\mathcal{B}^{\mathbb{C}}$ is controllable.*

Proof. This result follows immediately from the definitions of controllability and of the complex form $\mathcal{B}^{\mathbb{C}}$ of $\mathcal{B}$. □

In the commutative case there are many characterizations of controllability as can be seen in the following theorem.

Theorem 5.1.3. [54] *Let $R \in \mathbb{R}^{g\times r}[s, s^{-1}]$ be frr and $\mathcal{B} = \ker R$. Then the following conditions are equivalent:*

(*i*) *$\mathcal{B}$ is controllable;*

(*ii*) *R is left prime;*

(*iii*) *the Smith form of R is $\begin{bmatrix} I_g & 0 \end{bmatrix}$;*

(*iv*) $\operatorname{rank} R(\lambda, \lambda^{-1})$ *is constant for all* $0 \neq \lambda \in \mathbb{C}$*;*

(*v*) *there exists an image representation for $\mathcal{B}$, i.e., $\exists M \in \mathbb{R}^{r\times m}[s, s^{-1}]$ such that* $\mathcal{B} = \operatorname{im} M(\sigma, \sigma^{-1})$.

As we have shown in Example 3.1.3, if $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ is unimodular, then $U(\lambda, \lambda^{-1})$ is not necessarily invertible for all $0 \neq \lambda \in \mathbb{H}$. Clearly, in this case the unimodular matrix U is a kernel representation of the (trivially) controllable behavior $\mathcal{B} = \{0\}$, it is left prime and its quaternionic Smith form is I_g, but $\operatorname{rank} U(\lambda, \lambda^{-1})$ is not constant in $\mathbb{H}\backslash\{0\}$. Therefore the condition (iv) of Theorem 5.1.3, which corresponds to the Hautus criterion for state-space models, is not valid in the quaternionic case.

The following theorem shows that the other conditions of Theorem 5.1.3 remain valid in the quaternionic case. First we give an auxiliary result.

Lemma 5.1.4. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ and $\mathcal{B} = \ker R$.*

1. *Let $U \in \mathbb{H}^{r\times r}[s, s^{-1}]$ be unimodular. Then $\mathcal{B}$ is controllable if and only if $U(\mathcal{B})$ is controllable.*

2. *$\mathcal{B} = \mathcal{B}_1 \times (\mathbb{H}^{r_2})^{\mathbb{Z}}$ is controllable if and only if $\mathcal{B}_1$ is controllable.*

Proof. We will only consider the proof of the first statement since the second is trivially verified.

"If" part. Let $w_1, w_2 \in \mathcal{B}$. Then $v_l = Uw_l \in U(\mathcal{B})$, $l = 1, 2$. Since $U(\mathcal{B})$ is controllable there exists $\Delta > 0$, such that for $t_2 - t_1 > \Delta$, there exists a trajectory $v = Uw \in U(\mathcal{B})$ with the property

$$v(t) = \begin{cases} v_1(t), & t \leq t_1; \\ v_2(t), & t \geq t_2. \end{cases}$$

Suppose that $U^{-1} = U_H s^H + \cdots + U_L s^L \in \mathbb{H}^{g\times r}[s, s^{-1}]$ and let $\Delta' = \max\{|H|, |L|\}$. Then $U^{-1}v \in \mathcal{B}$ and is such that

$$\left(U^{-1}v\right)(t) = \begin{cases} \left(U^{-1}v_1\right)(t), & t \leq t_1 - \Delta'; \\ \left(U^{-1}v_2\right)(t), & t \geq t_2 + \Delta'. \end{cases}$$

But $U^{-1}v = w$ and $U^{-1}v_l = w_l$, $l = 1, 2$. Therefore we conclude that there exists $\widetilde{\Delta} = \Delta + 2\Delta' > 0$ such that for $\tau_2 - \tau_1 > \widetilde{\Delta}$, there exists $w \in \mathcal{B}$ with

$$w(t) = \begin{cases} w_1(t), & t \leq \tau_1; \\ w_2(t), & t \geq \tau_2. \end{cases}$$

and hence $\mathcal{B}$ is controllable.

"Only if" part. The proof is analogous to the last one, now with the roles of U and U^{-1} interchanged. $\square$

Theorem 5.1.5. *Let $R \in \mathbb{H}^{g\times r}[s,s^{-1}]$ be frr and $\mathcal{B} = \ker R(\sigma,\sigma^{-1})$. Then the following conditions are equivalent:*

(i) *$\mathcal{B}$ is controllable;*

(ii) *R is left prime;*

(iii) *the quaternionic Smith form of R is $\begin{bmatrix} I_g & 0 \end{bmatrix}$;*

(iv) *there exists an image representation, i.e., $\exists M \in \mathbb{H}^{r\times m}[s,s^{-1}]$ such that $\mathcal{B} =$* $\operatorname{im} M(\sigma,\sigma^{-1})$.

Proof. We will show that the implications $(ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i) \Rightarrow (ii)$ hold true.

$(ii) \Rightarrow (iii)$ Let $R = U\Gamma V$, with $U \in \mathbb{H}^{g\times g}[s,s^{-1}]$ and $V \in \mathbb{H}^{r\times r}[s,s^{-1}]$ unimodular, and

$$\Gamma = \left[\begin{array}{ccc|c} \gamma_1 & & & \\ & \ddots & & 0 \\ & & \gamma_g & \end{array}\right] \in \mathbb{H}^{g\times r}[s]$$

a quaternionic Smith form of R.

Partition V conformably as

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \; V_1 \in \mathbb{H}^{g\times r}[s,s^{-1}], \; V_2 \in \mathbb{H}^{(r-g)\times r}[s,s^{-1}],$$

then

$$R = WV_1,$$

where $W = U \operatorname{diag}(\gamma_1,\ldots,\gamma_g) \in \mathbb{H}^{g\times g}[s,s^{-1}]$.

The left primeness of R (see Definition 3.1.6) implies that W is unimodular, and therefore also $\operatorname{diag}(\gamma_1,\ldots,\gamma_g)$ is unimodular, which implies that the polynomials γ_l, $l = 1,\ldots,g$, are constant. Since, by definition, they are monic, we have that $\gamma_l = 1$, $l = 1,\ldots,g$, i.e., $\Gamma = \begin{bmatrix} I_g & 0 \end{bmatrix}$.

$(iii) \Rightarrow (iv)$ Let $w \in \mathcal{B}$, i.e., $Rw = 0$. Since the quaternionic Smith form of R is $\begin{bmatrix} I_g & 0 \end{bmatrix}$ there exist unimodular matrices $U \in \mathbb{H}^{g\times g}[s,s^{-1}]$ and $V \in \mathbb{H}^{r\times r}[s,s^{-1}]$ such that $URV = \begin{bmatrix} I_g & 0 \end{bmatrix}$. Then

$$Rw = 0 \Leftrightarrow URVV^{-1}w = 0 \Leftrightarrow \begin{cases} \begin{bmatrix} I_g & 0 \end{bmatrix} \widetilde{w} = 0 \\ w = V\widetilde{w} \end{cases}. \tag{5.2}$$

If $\widetilde{w} = \left[\begin{smallmatrix} \widetilde{w}_1 \\ \widetilde{w}_2 \end{smallmatrix}\right]$, with $\widetilde{w}_1 \in (\mathbb{H}^g)^{\mathbb{Z}}$, then (5.2) means that $\widetilde{w}_1 = 0$ and $\widetilde{w}_2$ is free. Partition V conformably as $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$. Then

$$w = V\widetilde{w} = V_1\widetilde{w}_1 + V_2\widetilde{w}_2 = V_2\widetilde{w}_2,$$

where $\widetilde{w}_2$ is free. Thus $\mathcal{B} \subseteq \operatorname{im} V_2$ and, since $RV_2 = 0$, $\operatorname{im} V_2 \subseteq \mathcal{B}$, therefore $\mathcal{B} = \operatorname{im} V_2$, i.e., $\mathcal{B}$ admits an image representation.

$(iv) \Rightarrow (i)$ Assume that $\mathcal{B} = \operatorname{im} M$ with $M = M_H s^H + \cdots + M_L s^L \in \mathbb{H}^{r\times m}[s, s^{-1}]$. Let $w_1, w_2 \in \mathcal{B}$. Then there exist $v_1, v_2 \in (\mathbb{H}^m)^{\mathbb{Z}}$ such that $w_1 = Mv_1$ and $w_2 = Mv_2$. Define $\Delta = H - L$. Given any $t_1 \in \mathbb{Z}$ let $t_2 > t_1 + \Delta$ and let $v \in (\mathbb{H}^m)^{\mathbb{Z}}$ be such that

$$v(t) = \begin{cases} v_1(t), & t \leq t_1 + H \\ v_2(t), & t > t_1 + H \end{cases}.$$

It is not difficult to check that $w = Mv$ is a trajectory in $\mathcal{B}$ such that

$$w(t) = \begin{cases} w_1(t), & t \leq t_1 \\ w_2(t), & t \geq t_2 \end{cases},$$

and hence $\mathcal{B}$ is controllable.

$(i) \Rightarrow (ii)$ Suppose that R is not left prime, i.e., $R = D\widetilde{R}$ where $\widetilde{R}$ is left prime and $D \in \mathbb{H}^{g\times g}[s, s^{-1}]$ is not unimodular. Since $\widetilde{R} = U \begin{bmatrix} I_g & 0 \end{bmatrix} V$, with U and V unimodular matrices of suitable dimensions, we have that

$$R = DU \begin{bmatrix} I_g & 0 \end{bmatrix} V = \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} V.$$

Let $w \in \mathcal{B}$. Then $Rw = 0$, which is equivalent to $\widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} Vw = 0$, i.e., $Vw \in \ker \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix}$. Therefore $\ker \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} \supset V(\mathcal{B})$. Conversely if $v \in \ker \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix}$, $w = V^{-1}v$ is such that $\widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} Vw = 0$, i.e., $Rw = 0$, and hence belongs to $\mathcal{B}$. Thus $\ker \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} \subset V(\mathcal{B})$ and we conclude that $\ker \widetilde{D} \begin{bmatrix} I_g & 0 \end{bmatrix} = V(\mathcal{B})$. By Lemma 5.1.4-1 we have that $\mathcal{B}$ is controllable if and only if $\ker \widetilde{D}\begin{bmatrix} I_g & 0 \end{bmatrix}$ is controllable, and by Lemma 5.1.4-2 this is equivalent to $\ker \widetilde{D}$ being controllable. Note that, by the same argument as above, $\widetilde{D}$ may be assumed to be a quaternionic Smith form. Let $\widetilde{w} \in \ker \widetilde{D}$, i.e., $\widetilde{D}\widetilde{w} = 0$. As $\widetilde{D}$ is not unimodular, we have that $\widetilde{D} \neq I$. Therefore $\widetilde{D} = \operatorname{diag}(d_1, \ldots, d_g)$, where, without loss of generality, $d_1(s) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_0$ with $\alpha_0 \neq 0$, $n \geq 1$. Thus

$$\widetilde{D}\widetilde{w} = 0 \Leftrightarrow \begin{cases} d_1\widetilde{w}_1 & = & 0 \\ d_2\widetilde{w}_2 & = & 0 \\ \vdots & & \\ d_g\widetilde{w}_g & = & 0 \end{cases},$$

and the solutions of $d_1\widetilde{w}_1 = 0$ are uniquely determined by the values of $\widetilde{w}_1$ at the points $t, t+1, \ldots, t+n-1$ (initial conditions). Therefore, if $\widetilde{w}_1 \in \ker \widetilde{D}$ has nonzero initial conditions then it is not concatenable in $\mathcal{B}$ with the zero trajectory. So, we conclude that $\ker \widetilde{D}$ is not controllable and hence neither is $\mathcal{B}$. $\square$

Remark 5.1.6. Lemma 5.1.2 says that it is possible to check whether a quaternionic behavior is controllable by analysing its complex form. However, in general, this method, besides increasing the size of the matrices involved and hence the computational complexity, may also transform the problem into a less intuitive one. For instance, let $R = \begin{bmatrix} s+\mathbf{j} & (s+\mathbf{j})\mathbf{i} \end{bmatrix}$ and $\mathcal{B} = \ker R$. It is possible to conclude immediately that $\mathcal{B}$ is not controllable since $\begin{bmatrix} s+\mathbf{j} & 0 \end{bmatrix}$ is obviously a quaternionic Smith form of R. On the other hand, looking at the corresponding complex adjoint matrix $R^c = \left[\begin{smallmatrix} s & \mathbf{i}s & 1 & -\mathbf{i} \\ -1 & -\mathbf{i} & s & -\mathbf{i}s \end{smallmatrix}\right]$ it is not so evident whether $\mathcal{B}^{\mathbb{C}}$ is controllable or not. $\square$

As happens in the real or complex case, we define the *controllable subbehavior* of a given behavior $\mathcal{B}$, denoted by $\mathcal{B}_c$, as the largest controllable subbehavior of $\mathcal{B}$, in the sense that it contains all the controllable subbehaviors of $\mathcal{B}$. In the sequel we show how to obtain $\mathcal{B}_c$ from any kernel representation of $\mathcal{B}$.

Let then $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ be frr and such that $\mathcal{B} = \ker R(\sigma, \sigma^{-1})$. Note that the matrix R can be factorized as $R = FC$, where $F \in \mathbb{H}^{g\times g}[s, s^{-1}]$ has full rank and $C \in \mathbb{H}^{g\times r}[s, s^{-1}]$ is left prime. Indeed, letting $R = U\begin{bmatrix} \Gamma & 0 \end{bmatrix}V = U\Gamma\begin{bmatrix} I_g & 0 \end{bmatrix}V$ where $\begin{bmatrix} \Gamma & 0 \end{bmatrix}$

is a quaternionic Smith form of R and U and V are unimodular matrices, we can take $F = U\Gamma$ and $C = \begin{bmatrix} I_g & 0 \end{bmatrix} V$. We will show that $\mathcal{B}_c = \ker C$. We start by showing that $\ker C$ does not depend on the kernel representation of $\mathcal{B}$.

Let $\mathcal{B} = \ker R_1 = \ker R_2$ with $R_1, R_2 \in \mathbb{H}^{g\times r}[s, s^{-1}]$ frr and factorize

$$R_1 = F_1 C_1 \quad \text{and} \quad R_2 = F_2 C_2, \tag{5.3}$$

where $F_1, F_2 \in \mathbb{H}^{g\times g}[s, s^{-1}]$ have full rank and $C_1, C_2 \in \mathbb{H}^{g\times r}[s, s^{-1}]$ are left prime. We claim that $\ker C_1 = \ker C_2$. By Corollary 4.2.11 there exits a unimodular matrix $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ such that $R_1 = UR_2$ and then, from equation (5.3) we obtain

$$UF_2C_2 = F_1C_1. \tag{5.4}$$

Since F_1 has full rank it admits a rational inverse, i.e., $F_1^{-1} \in \mathbb{H}^{g\times g}(s, s^{-1})$. Moreover, as stated in [39], any left prime matrix admits a L-polynomial right inverse and consequently there exists $C_2^\sharp \in \mathbb{H}^{r\times g}[s, s^{-1}]$ such that $C_2C_2^\sharp = I_g$. Then equation 5.4 can be written as

$$F_1^{-1}UF_2C_2 = C_1 \tag{5.5}$$

$$\Leftrightarrow \quad F_1^{-1}UF_2 = C_1C_2^\sharp. \tag{5.6}$$

From equation (5.6) we conclude that $F_1^{-1}UF_2$ is a L-polynomial matrix and by (5.5) it must be unimodular since it is a square left factor of the left prime matrix C_1. Hence, by Corollary 4.2.11 and equation (5.5) we have that $\ker C_1 = \ker C_2$ as claimed.

We are now in a position to show that $\ker C$, with C defined as above, is indeed the largest controllable subbehavior of $\mathcal{B} = \ker R$, i.e., that $\mathcal{B}_c = \ker C$. Let $\widetilde{\mathcal{B}} = \ker \widetilde{R}$ be a controllable subbehavior of $\mathcal{B}$. Since $\widetilde{\mathcal{B}} \subset \mathcal{B}$, by Theorem 4.2.10, there exists a polynomial matrix X such that $R = X\widetilde{R}$. Note that X has frr because so has R. By Theorem 5.1.5, $\widetilde{R}$ is left prime. If X is a square matrix then, as proved before, $\widetilde{\mathcal{B}} = \ker \widetilde{R} = \ker C$. If not, factorize $X = \widetilde{F}\widetilde{C}$, where $\widetilde{F}$ is a square full rank matrix and $\widetilde{C}$ is left prime. Note that the matrix $\widetilde{C}\widetilde{R}$ is left prime and, since $R = \widetilde{F}\widetilde{C}\widetilde{R}$, we have again that $\ker \widetilde{C}\widetilde{R} = \ker C$. By Theorem 4.2.10, $\widetilde{\mathcal{B}} \subset \ker C$ and hence $\ker C$ is the largest controllable subbehavior of $\mathcal{B}$, i.e. $\mathcal{B}_c = \ker C$.

5.2 Autonomy

Autonomous behaviors are roughly speaking the ones whose trajectories are completely determined once their 'past' is known and can therefore be considered the extreme opposite of controllable behaviors. After characterizing autonomy, we will show that any kernel behavior can be written as a direct sum of a controllable and an autonomous subbehavior.

Definition 5.2.1. [43] A behavior $\mathcal{B} \subseteq (\mathbb{H}^q)^{\mathbb{Z}}$ is called *autonomous* if for all $w_1, w_2 \in \mathcal{B}$ and all $t_0 \in \mathbb{Z}$

$$w_1(t) = w_2(t) \text{ for } t \leq t_0 \;\Rightarrow\; w_1 \equiv w_2.$$

Clearly, if $\mathcal{B}$ is shift-invariant, then t_0 can be replaced by 0 in the previous definition. Moreover, if $\mathcal{B}$ is a behavior linear on the right, then $\mathcal{B}$ is autonomous if and only if $w(t) = 0,\ t \leq 0$, implies that $w(t) = 0$ for every t. As in the commutative case the following proposition holds.

Proposition 5.2.2. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ and $\mathcal{B} = \ker R(\sigma, \sigma^{-1})$. Then the following conditions are equivalent:*

(i) $\mathcal{B}$ is autonomous;

(ii) R has full column rank;

(iii) $\mathcal{B}$ is a finite dimensional right-vector space.

Proof. We will show that the equivalences $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$ hold true.

$(i) \Rightarrow (ii)$ Suppose that R has not full column rank. Then there exist unimodular matrices $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ and $V = V_N s^N + \cdots + V_0 \in \mathbb{H}^{r\times r}[s]$ such that

$$\widetilde{R} = URV = \begin{bmatrix} \widehat{R} & | & 0 \end{bmatrix},\ \widehat{R} \in \mathbb{H}^{g\times(r-1)}[s, s^{-1}].$$

Let $\widetilde{w} = \begin{bmatrix} 0 & \cdots & 0 & \widetilde{w}_r \end{bmatrix}^T \in (\mathbb{H}^r)^{\mathbb{Z}}$ with

$$\widetilde{w}_r(t) = \begin{cases} 0, & t \leq N \\ 1, & t > N \end{cases}.$$

Clearly $\widetilde{R}(\sigma,\sigma^{-1})\widetilde{w} = 0$. Let $w = V\widetilde{w}$. Note that $w(t) = 0,\ t \leq 0$. Letting t^* be sufficiently big,

$$\begin{aligned} w(t^*) \;=\; \big(V(\sigma)\widetilde{w}\big)(t^*) \;&=\; \Big(\big(V_N\sigma^N + \cdots + V_1\sigma + V_0\big)\widetilde{w}\Big)(t^*) \\ &=\; \big(\underbrace{V_N + \cdots V_1 + V_0}_{V(1)}\big)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

The unimodularity of V implies that $V(1)$ is invertible (cf page 53) and thus $w(t^*) \neq 0$. Moreover

$$R(\sigma,\sigma^{-1})w = U^{-1}(\sigma,\sigma^{-1})\widetilde{R}(\sigma,\sigma^{-1})\widetilde{w} = 0,$$

and therefore $w \in \mathcal{B}$. This implies that $\mathcal{B}$ is not autonomous.

(ii) $\Rightarrow$ *(i)* Suppose that R has full column rank. Then, by the procedures of section 4.3, we know that the solutions of $R(\sigma,\sigma^{-1})w = 0$ are uniquely determined by the values of w in a finite number of consecutive points (cf Section 4.3.2) and therefore $\mathcal{B}$ is autonomous.

(ii) $\Rightarrow$ *(iii)* Suppose that R has full column rank. Then there exist unimodular matrices U and V such that the quaternionic Smith form of R is

$$\Gamma(s) = \operatorname{diag}\big(\gamma_1(s),\ldots,\gamma_r(s)\big)$$

with $\gamma_l(s) \neq 0,\ l = 1,\ldots,r$. Let $w \in \mathcal{B}$, i.e., $R(\sigma,\sigma^{-1})w = 0$. Let $\widetilde{w} = V^{-1}w$. Then,

$$R(\sigma,\sigma^{-1})w = \Gamma(\sigma)\widetilde{w} = 0,$$

and therefore $\widetilde{\mathcal{B}} = V^{-1}(\mathcal{B}) = \ker\Gamma$. It can easily be seen that $\ker\Gamma$ is a finite dimensional vector space. Hence the same happens with $\mathcal{B} = V(\widetilde{\mathcal{B}})$.

(iii) $\Rightarrow$ *(ii)* Suppose that R has not full column rank and let $w \in \mathcal{B}$. By the arguments presented in Section 4.2.3, there exists a permutation matrix Π such that $w = \Pi\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$, where $u \in \left(\mathbb{H}^{r-g}\right)^{\mathbb{T}}$ can be chosen freely. Since $\left(\mathbb{H}^{r-g}\right)^{\mathbb{T}}$ is infinite dimensional, also is $\mathcal{B}$. $\square$

As in the commutative case every behavior can be decomposed into a controllable and an autonomous part. The proof is analogous to the one of [43, Theorem 5.2.14].

Theorem 5.2.3. *Every quaternionic behavior $\mathcal{B}$ can be decomposed as*

$$\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a,$$

where $\mathcal{B}_c$ and $\mathcal{B}_a$ are, respectively, the controllable and an autonomous subbehavior of $\mathcal{B}$.

Proof. Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ be full row rank and $\mathcal{B} = \ker R$. Then there exist unimodular matrices $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ and $V \in \mathbb{H}^{r\times r}[s]$ such that the quaternionic Smith form of R is

$$URV = \begin{bmatrix}\Gamma & | & 0\end{bmatrix}, \quad \Gamma \in \mathbb{H}^{g\times g}[s].$$

Partition $\widetilde{w} = \left[\begin{smallmatrix}\widetilde{w}_1\\ \widetilde{w}_2\end{smallmatrix}\right] \in (\mathbb{H}^g)^{\mathbb{Z}}$ conformably and define the behavior

$$\widetilde{\mathcal{B}} = V^{-1}(\mathcal{B}) = \left\{\widetilde{w} \in (\mathbb{H}^g)^{\mathbb{Z}} \mid \Gamma(\sigma)\widetilde{w}_1 = 0,\ \widetilde{w}_2 \text{ is free}\right\}. \tag{5.7}$$

If $\Gamma(s)$ is the identity matrix then, by Theorem 5.1.5 the behavior $\mathcal{B}$ is controllable and the result follows with $\mathcal{B}_c = \mathcal{B}$ and $\mathcal{B}_a = \{0\}$.

Suppose that $\Gamma(s) \neq I$. Then it is not difficult to see that the behavior $\widetilde{\mathcal{B}}$ can be decomposed as $\widetilde{\mathcal{B}} = \widetilde{\mathcal{B}}_c \oplus \widetilde{\mathcal{B}}_a$ where

$$\widetilde{\mathcal{B}}_c = \left\{\widetilde{w} \in (\mathbb{H}^g)^{\mathbb{Z}} \mid \widetilde{w}_1 = 0,\ \widetilde{w}_2 \text{ is free}\right\} = \ker \widetilde{R}_c, \quad \widetilde{R}_c = \begin{bmatrix} I & 0\end{bmatrix},$$

and

$$\widetilde{\mathcal{B}}_a = \left\{\widetilde{w} \in (\mathbb{H}^g)^{\mathbb{Z}} \mid \Gamma(\sigma)\widetilde{w}_1 = 0,\ \widetilde{w}_2 = 0\right\} = \ker \widetilde{R}_a, \quad \widetilde{R}_a = \begin{bmatrix} \Gamma & 0\\ 0 & I\end{bmatrix}.$$

By Theorem 5.1.5 we have that the behavior $\widetilde{\mathcal{B}}_c$ is controllable and, since the matrix $\widetilde{R}_a$ has full column rank, by Proposition 5.2.2, the behavior $\widetilde{B}_a$ is autonomous. By (5.7), $\mathcal{B} = V(\widetilde{\mathcal{B}})$, i.e., $\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a$, where $\mathcal{B}_c = V(\widetilde{\mathcal{B}}_c) = \ker \widetilde{R}_c V^{-1}$ and $\mathcal{B}_a = V(\widetilde{\mathcal{B}}_a) = \ker \widetilde{R}_a V^{-1}$. Note that, by construction, $\mathcal{B}_c$ is indeed the controllable subbehavior of $\mathcal{B}$ and that $\mathcal{B}_a$ is autonomous since the matrix $\widetilde{R}_a V^{-1}$ has full column rank. $\square$

5.3 Stability

Stability is an important property of dynamical systems. Loosely speaking, a dynamical system is said to be stable if small perturbations produce small effects. In this section simple and asymptotic stability for a generic quaternionic behavior are defined and characterized.

Definition 5.3.1. [43, Def. 7.2.1] A linear on the right behavior $\mathcal{B}$ is said to be *stable* if for every $w \in \mathcal{B}$, $||w(t)||$ is bounded for all $t > 0$. If, in addition, $\lim_{t\to+\infty} w(t) = 0$, the behavior is *asymptotically stable.*

Remark 5.3.2. Note that this definition implies that the trajectories of a stable behavior cannot contain free components and therefore a stable behavior only admits full column rank kernel representations. Then, by Proposition 5.2.2, we have that every stable behavior is autonomous. □

Lemma 5.3.3. *A quaternionic behavior $\mathcal{B}$ is (asymptotically) stable if and only if its complex form, $\mathcal{B}^{\mathbb{C}}$, is (asymptotically) stable.*

Proof. Let $w(t) = w_1(t) + w_2(t)\mathbf{j}$. By the definition of $\mathcal{B}^{\mathbb{C}}$ we have that $w \in \mathcal{B}$ if and only if $w^{\mathbb{C}} = \left[\begin{smallmatrix} w_1 \\ -\overline{w}_2 \end{smallmatrix}\right] \in \mathcal{B}^{\mathbb{C}}$. Clearly, $||w(t)||$ is bounded if and only if $||w^{\mathbb{C}}(t)||$ is bounded and, moreover, $\lim_{t\to+\infty} w(t) = 0$ if and only if $\lim_{t\to+\infty} w_1(t) = \lim_{t\to+\infty} w_2(t) = 0$. Therefore, $\mathcal{B}$ is (asymptotically) stable if and only if $\mathcal{B}^{\mathbb{C}}$ is (asymptotically) stable. □

The characterization of asymptotically stable real or complex behaviors defined by $\mathcal{B} = \ker R$ can be given in terms of the roots of $\det R$ as follows.

Theorem 5.3.4. [43] *Let $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ be a behavior given as $\mathcal{B} = \ker R$, with $R \in \mathbb{C}^{r\times r}[s, s^{-1}]$. Then $\mathcal{B}$ is asymptotically stable if and only if all the roots of $\det R(s)$ have norm less than one.*

With the definition of the polynomial determinant Pdet for quaternionic polynomial matrices given in Section 3.2 it is possible to extend Theorem 5.3.4 to the quaternionic case.

Consider first the *stability region*

$$\mathcal{S}_{\mathbb{Z}} = \{q \in \mathbb{H} : \ |q| < 1\},$$

which extends to the quaternionic case the usual complex stability regions used for discrete-time real systems. Note that, by Proposition 1.1.2, $\mathcal{S}_{\mathbb{Z}}$ satisfies

$$\lambda \in \mathcal{S}_{\mathbb{Z}} \Rightarrow [\lambda] \subseteq \mathcal{S}_{\mathbb{Z}}. \tag{5.8}$$

We start by considering the first order quaternionic system

$$w(t+1) = Aw(t) \tag{5.9}$$

with $A \in \mathbb{H}^{n\times n}$. As shown in Section 4.3.1, the solutions of (5.9) are given by

$$w(t) = VJ^t q, \tag{5.10}$$

where $V \in \mathbb{H}^{n\times n}$, J is the Jordan form of A and $q \in \mathbb{H}$ is an initial condition.

Taking into account the special structure of the Jordan form, it is possible to prove that the components of $\widetilde{w}(t) = V^{-1}w(t) = J^t q$ are given by

$$\lambda^t p(t)$$

where the λ's are the elements in the diagonal of J and $p(t)$ is a suitable quaternionic polynomial.

On the other hand, if λ is a diagonal element of J, there exists a suitable initial value $q = \mathbf{e}_r$ (where $\mathbf{e}_r$ is the r-th vector of the canonical basis of $\mathbb{R}^n \subset \mathbb{H}^n$) such that

$$\widetilde{w}(t) = \lambda^t \mathbf{e}_r$$

is a solution of (5.9).

It turns out that the elements in the diagonal of J correspond to the standard right eigenvalues of A. Together with Theorem 3.2.7, this allows to characterize the asymptotic stability of $w(t+1) = Aw(t)$ in terms of the right spectrum $\sigma_r(A)$ of A, or equivalently, in terms of the zeros of $\mathrm{Pdet}(sI - A)$.

Proposition 5.3.5. *Let $A \in \mathbb{H}^{n\times n}$. Then the following statements are equivalent.*

(i) The quaternionic system described by $w(t+1) = Aw(t)$ is asymptotically stable.

(ii) $\sigma_r(A) \subset \mathcal{S}_{\mathbb{Z}}$

(iii) All the zeros of $\operatorname{Pdet}(sI - A)$ *lie in $S_{\mathbb{Z}}$.*

Proof. The equivalence between (ii) and (iii) is a direct consequence of Theorem 3.2.7.

$(i) \Rightarrow (ii)$ If $\sigma_r(A) \not\subset S_{\mathbb{Z}}$, there exists a standard right eigenvalue λ of A such that $|\lambda| \geq 1$. Thus, keeping the notation of the previous considerations, for a suitable r, $w(t) = V\lambda^t \mathbf{e}_r$, is a solution of $w(t+1) = Aw(t)$. Since, obviously, $\lim\limits_{t\to+\infty} w(t) \neq 0$ the system is not asymptotically stable.

$(ii) \Rightarrow (i)$ If $\sigma_r(A) \subset S_{\mathbb{Z}}$, then $\lim\limits_{t\to+\infty} \widetilde{w}(t) = \lim\limits_{t\to+\infty} J^t q = 0$, for all $q \in \mathbb{H}^n$. Since $w(t) = V\widetilde{w}(t)$, this clearly implies that $\lim\limits_{t\to+\infty} w(t) = 0$, for every solution $w(t)$ of $w(t+1) = Aw(t)$, proving that the system is asymptotically stable. □

Remark 5.3.6. The characterization of asymptotic stability of the first order continuous-time quaternionic system $\dot{w}(t) = Aw(t)$ is analogous. Note that, in this case, the stability region is

$$\mathcal{S}_{\mathbb{R}} = \{q \in \mathbb{H} : \operatorname{Re} q < 0\}$$

which, by Proposition 1.1.2, also satisfies

$$\lambda \in \mathcal{S}_{\mathbb{R}} \Rightarrow [\lambda] \subseteq \mathcal{S}_{\mathbb{R}}.$$

□

Consider now a quaternionic system described by a higher order matrix differential equation

$$R(\sigma)\, w = 0 \tag{5.11}$$

with $R(s) = R_m s^m + \cdots + R_1 s + R_0 \in \mathbb{H}^{r\times r}[s]$. Assume first that $R(s)$ has full rank. As it was shown in Section 4.3.2, there exists a unimodular matrix $U(s) \in \mathbb{H}^{r\times r}[s]$ such that

$$U(s)R(s) = I_r s^m + \widetilde{R}_{m-1} s^{m-1} + \cdots + \widetilde{R}_1 s + \widetilde{R}_0 = \widetilde{R}(s),$$

and defining $x(t) = \begin{bmatrix} w(t)^T & w(t+1)^T & \cdots & w(t+m-1)^T \end{bmatrix}^T$, we obtain the alternative system description

$$\begin{cases} x(t+1) = Ax(t) \\ w(t) = Cx(t) \end{cases} \tag{5.12}$$

with

$$A = \begin{bmatrix} 0 & I_r & 0 & \cdots & 0 \\ 0 & 0 & I_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_r \\ -\widetilde{R}_0 & -\widetilde{R}_1 & -\widetilde{R}_2 & \cdots & -\widetilde{R}_{m-1} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} I_r & 0 & \cdots & 0 \end{bmatrix}.$$

Note that

$$\operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{l-1} \end{bmatrix} = l, \quad \text{with } l = rm.$$

This is equivalent to say that $\begin{bmatrix} sI - A \\ C \end{bmatrix}$ has a left inverse $W(s)$ (see [43, Theorem 5.3.15]). Therefore, it follows from (5.12) that

$$x(t) = W(\sigma) \begin{bmatrix} 0 \\ I \end{bmatrix} w(t).$$

Together with the second equation of (5.12) this implies that $\lim_{t \to +\infty} w(t) = 0$ if and only if $\lim_{t \to +\infty} x(t) = 0$. Hence the asymptotical stability of the original system is equivalent to the asymptotical stability of (5.12), which in turn, by Proposition 5.3.5 and Corollary 3.2.9, is equivalent to say that all the zeros of $\operatorname{Pdet} R(s)$ lie in $\mathcal{S}_{\mathbb{Z}}$.

If $R(s)$ has not full rank, $\operatorname{Pdet} R(s) \equiv 0$. In this case, by Proposition 5.2.2 the corresponding system is not autonomous and hence, by Remark 5.3.2, is not asymptotically stable.

We conclude in this way that asymptotic stability of (5.11) can be characterized in terms of the zeros of $\operatorname{Pdet} R(s)$ as follows.

Theorem 5.3.7. *Let $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ be a quaternionic behavior with kernel representation $R \in \mathbb{H}^{r\times r}[s]$. This system is asymptotically stable if and only if all the zeros of* $\operatorname{Pdet} R$ *lie in $\mathcal{S}_{\mathbb{Z}}$.*

In the commutative case, it is well-known [28] that given a full rank matrix $R \in \mathbb{C}^{n\times n}[s]$, λ is a root of $\det R(s)$ if and only $\operatorname{rank} R(\lambda) < n$. Thus, Theorem 5.3.4 can be written in the following alternative form.

Theorem 5.3.8. *Let $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ be a behavior given as $\mathcal{B} = \ker R$, with $R \in \mathbb{C}^{r\times r}[s]$. Then $\mathcal{B}$ is asymptotically stable if and only if* $\operatorname{rank} R(\lambda) = r$ *for all λ such that $|\lambda| < 1$.*

Note that this characterization is given in terms of evaluation of polynomials, which, as mentioned in Section 2.1, in the quaternionic case is not a ring homomorphism. This fact suggests that for quaternionic behaviors such a result does not hold. Indeed, consider the quaternionic matrix $U \in \mathbb{H}^{2\times 2}[s]$ given in Example 3.1.3 and let $\mathcal{B} = \ker U$. Since U is unimodular it is clear that $\mathcal{B} = \{0\}$ and hence it is asymptotically stable. But, $\left|1 + \frac{1}{2}\mathbf{j}\right| = \sqrt{5}/2 \geq 1$ and, as it was shown in the example, the matrix $U\left(1 + \frac{1}{2}\mathbf{j}\right)$ is not invertible, i.e., $\operatorname{rank} U\left(1 + \frac{1}{2}\mathbf{j}\right) < 2$.

As we have seen, the polynomial determinant Pdet provides a criterion for asymptotic stability, but it can not discriminate between stable or unstable behaviors in case of zeros with norm equal to one. This can be seen in the following example.

Example 5.3.9. Let

$$R_1(s) = \begin{bmatrix} s-1 & 0 \\ 0 & s-1 \end{bmatrix} \qquad \text{and} \qquad R_2(s) = \begin{bmatrix} s-1 & 1 \\ 0 & s-1 \end{bmatrix}.$$

It follows that

$$\mathcal{B}_1 = \ker R_1(\sigma) = \left\{ \begin{bmatrix} q_1 \\ q_2 \end{bmatrix},\ q_1, q_2 \in \mathbb{H} \right\}$$

and

$$\mathcal{B}_2 = \ker R_2(\sigma) = \left\{ \begin{bmatrix} -tc_2 + c_1 \\ c_2 \end{bmatrix},\ c_1, c_2 \in \mathbb{H} \right\}.$$

Then it is clear that $\mathcal{B}_1$ is stable while $\mathcal{B}_2$ is not stable but, by definition, $\mathrm{Pdet}(R_1) = \mathrm{Pdet}(R_2) = (s-1)^4$. $\square$

As became clear in Section 2.2.2, the issue of the multiplicity of the zeros of a quaternionic polynomial is non trivial and so far we were not able to provide a characterization of stability in terms of the polynomial determinant Pdet.

However, an alternative characterization based on the quaternionic Smith form can be given both for the asymptotic stability and the stability of a quaternionic behavior. Before, we introduce the necessary terminology and preliminary results.

The notion of stable polynomial is generalized to the quaternionic case in the following definition, where $\overline{\mathcal{S}_{\mathbb{Z}}}$ denotes the closure of $\mathcal{S}_{\mathbb{Z}}$ and, by definition of multiplicity, $\mu_\lambda(p) > 0 \Leftrightarrow p(\lambda) = 0$.

Definition 5.3.10. When dealing with discrete-time systems, $p \in \mathbb{H}[s]$ is

- asymptotically stable in $\mathbb{X} \subseteq \mathbb{H}$ if, for any $\lambda \in \mathbb{X}$, $\mu_\lambda(p) > 0 \Rightarrow \lambda \in \mathcal{S}_{\mathbb{Z}}$;
- stable in $\mathbb{X} \subseteq \mathbb{H}$ if, for any $\lambda \in \mathbb{X}$, $\mu_\lambda(p) > 0 \Rightarrow \lambda \in \overline{\mathcal{S}_{\mathbb{Z}}}$ and $\mu_\lambda(p) > 1 \Rightarrow \lambda \in \mathcal{S}_{\mathbb{Z}}$.

In what follows we do not specify $\mathbb{X}$ when $\mathbb{X} = \mathbb{H}$.

Lemma 5.3.11. *The polynomial $p \in \mathbb{H}[s]$ is (asymptotically) stable if and only if $\mathcal{M}_p$ is (asymptotically) stable in $\mathbb{C}$.*

Proof. First, let us prove that

$$(\mu_\lambda(p) > l \;\Rightarrow\; \lambda \in \mathcal{S}_{\mathbb{Z}}) \;\Leftrightarrow\; (\mu_{[\lambda]}(p) > l \;\Rightarrow\; \lambda \in \mathcal{S}_{\mathbb{Z}}).$$

To show the implication "$\Rightarrow$", assume that $(\mu_\lambda(p) > l \;\Rightarrow\; \lambda \in \mathcal{S}_{\mathbb{Z}})$ and suppose that $\mu_{[\lambda]}(p) > l$. Then, by definition, there exists ν such that $\lambda \in [\nu]$ and $\mu_\nu(p) > l$ and so, by hypothesis, $\nu \in \mathcal{S}_{\mathbb{Z}}$. Hence, by condition (5.8), $\lambda \in \mathcal{S}_{\mathbb{Z}}$.

As for the implication "$\Leftarrow$", note that $\mu_{[\lambda]}(p) \geq \mu_\lambda(p)$. Thus, assuming that $(\mu_{[\lambda]}(p) > l \;\Rightarrow\; \lambda \in \mathcal{S}_{\mathbb{Z}})$, if $\mu_\lambda(p) > l$ then $\mu_{[\lambda]}(p) > l$ and so $\lambda \in \mathcal{S}_{\mathbb{Z}}$.

Finally, by Proposition 1.1.3 there always exists $\nu \in [\lambda] \cap \mathbb{C}$ and, by Proposition 2.2.32, $\mu_{[\lambda]}(p) = \mu_\nu\big(\mathcal{M}_p\big)$. The result then follows since we showed that in Definition 5.3.10 every condition on p with $\lambda \in \mathbb{H}$ can be equivalently written in terms of $\mathcal{M}_p$ with $\lambda \in \mathbb{C}$. $\square$

Theorem 5.3.12. *If $\Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_r)$ is a quaternionic Smith form of a kernel representation of the behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$, this behavior is (asymptotically) stable if and only if γ_r is (asymptotically) stable.*

Proof. By Lemma 5.3.3, stability of $\mathcal{B}$ is equivalent to stability of its complex form $\mathcal{B}^{\mathbb{C}}$, and, by Proposition 4.2.6, if $\mathcal{B} = \ker R$ then $\mathcal{B}^{\mathbb{C}} = \ker R^c$. Thus, by [43, Thm. 7.2.2] and by Theorem 3.4.14, $\mathcal{B}^{\mathbb{C}}$ is (asymptotically) stable if and only if $\mathcal{M}_{\gamma_r}$ is (asymptotically) stable in $\mathbb{C}$. The result is then a consequence of Lemma 5.3.11. $\square$

5.4 Stabilizability

A property which is weaker than controllability is stabilizability. In a stabilizable behavior, we may steer asymptotically, i.e., in infinite time, any trajectory to zero. Therefore, in a certain sense, stabilizability may be regarded as asymptotic controllability.

Definition 5.4.1. [43] A dynamical system with behavior $\mathcal{B}$ is called *stabilizable* if for every trajectory $w \in \mathcal{B}$ there exists a trajectory $w' \in \mathcal{B}$ such that

$$w'(t) = w(t)\ ,\ t \leq 0, \qquad \text{and} \qquad \lim_{t \to +\infty} w'(t) = 0.$$

By Theorem 5.2.3 we know that every behavior $\mathcal{B}$ can be written as a direct sum $\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a$, where $\mathcal{B}_c$ and B_a are, respectively, the controllable and an autonomous subbehavior of $\mathcal{B}$. The next theorem states the equivalence between the stabilizability of $\mathcal{B}$ and the asymptotic stability of $\mathcal{B}_a$.

Theorem 5.4.2. *Given a behavior $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$, the following statements are equivalent:*

(i) *$\mathcal{B}$ is stabilizable.*

(ii) *$\mathcal{B}$ allows a controllable-autonomous decomposition $\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a$ with $\mathcal{B}_a$ asymptotically stable.*

Proof. *(ii)* $\Rightarrow$ *(i)* Let $\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a$, with $\mathcal{B}_a$ asymptotically stable, and $w \in \mathcal{B}$. Then $w = w_a + w_c$ where $w_a \in \mathcal{B}_a$ and $w_c \in \mathcal{B}_c$. By the controllability of $\mathcal{B}_c$, there exists $t_1 \in \mathbb{Z}$ such that

$$w_c'(t) = \begin{cases} w_c(t), & t \leq 0 \\ 0 & t > t_1 \end{cases} \in \mathcal{B}_c.$$

Let $w' = w_a + w_c' \in \mathcal{B}$. Note that $w'(t) = w(t)$ for $t \leq 0$. Moreover, by hypothesis, $\mathcal{B}_a$ is asymptotically stable and therefore $\lim_{t\to+\infty} w_a(t) = 0$, implying thus that $\lim_{t\to+\infty} w'(t) = 0$. Hence $\mathcal{B}$ is stabilizable.

(i) $\Rightarrow$ *(ii)* Assume that $\mathcal{B}$ is stabilizable and has a frr representation R. Similar to what was done in Theorem 5.2.3, let U and V be unimodular matrices such that

$$URV = \begin{bmatrix} \Gamma & 0 \end{bmatrix}$$

is the quaternionic Smith form of R. Then $\mathcal{B} = V(\widetilde{\mathcal{B}})$ with

$$\widetilde{\mathcal{B}} = \left\{ \widetilde{w} = \left[\begin{smallmatrix} \widetilde{w}_1 \\ \widetilde{w}_2 \end{smallmatrix}\right] \in (\mathbb{H}^r)^{\mathbb{Z}} \mid \Gamma(\sigma)\widetilde{w}_1 = 0,\ \widetilde{w}_2 \text{ is free} \right\}.$$

Clearly, $\mathcal{B}$ is stabilizable if and only if $\widetilde{\mathcal{B}}$ is. But $\widetilde{\mathcal{B}} = \widetilde{\mathcal{B}}_c \oplus \widetilde{\mathcal{B}}_a$, with

$$\widetilde{\mathcal{B}}_c = \ker \begin{bmatrix} I & 0 \end{bmatrix} \quad \text{and} \quad \widetilde{\mathcal{B}}_a = \ker \begin{bmatrix} \Gamma & 0 \\ 0 & I \end{bmatrix},$$

and therefore it is a stabilizable behavior if and only if $\widetilde{\mathcal{B}}_a$ is asymptotically stable.

Taking into account the proof of Theorem 5.2.3, $\mathcal{B}$ allows the controllable-autonomous decomposition $\mathcal{B} = \mathcal{B}_c \oplus \mathcal{B}_a$, with $\mathcal{B}_c = V(\widetilde{\mathcal{B}}_c)$ and $\mathcal{B}_a = V(\widetilde{\mathcal{B}}_a)$. Since $\widetilde{\mathcal{B}}_a$ is asymptotically stable, so is $\mathcal{B}_a$, yielding the desired result. □

As happens for stable behaviors, the following result holds.

Lemma 5.4.3. *A quaternionic behavior $\mathcal{B}$ is stabilizable if and only if its complex form, $\mathcal{B}^{\mathbb{C}}$, is stabilizable.*

The characterization of stabilizability for a complex behavior $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ is given by the next result, which is the discrete version of [43, Thm. 5.2.30].

Theorem 5.4.4. *Let $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ be a complex behavior with kernel representation $R \in \mathbb{C}^{g\times r}[s]$. Then $\mathcal{B}$ is stabilizable if and only if* $\operatorname{rank} R(\lambda)$ *is constant for all $\lambda \in \mathbb{C}$ such that $|\lambda| \geq 1$.*

Note that, as happened with Theorem 5.3.8, it is not possible to generalize Theorem 5.4.4 for the quaternionic case. An alternative characterization of stabilizability of quaternionic behaviors given in terms of the quaternionic Smith form is the following.

Theorem 5.4.5. *Let $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ be a quaternionic behavior with kernel representation $R \in \mathbb{H}^{g\times r}[s]$ and let $\Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ be a quaternionic Smith form of R. Then*

$$\mathcal{B} \textit{ is stabilizable } \Leftrightarrow \big(\gamma_n(\lambda) = 0 \Rightarrow |\lambda| < 1,\ \lambda \in \mathbb{H}\big).$$

Proof. As mentioned in Lemma 5.4.3, $\mathcal{B}$ is stabilizable if and only if $\mathcal{B}^{\mathbb{C}}$ is stabilizable. To check this property, we may analyze the complex Smith form of R^c,

$$\Delta = \operatorname{diag}\left(\mathcal{F}_{\gamma_1}, \mathcal{M}_{\gamma_1}, \dots, \mathcal{F}_{\gamma_n}, \mathcal{M}_{\gamma_n}\right) \in \mathbb{R}^{2g\times 2r}[s].$$

Since R^c and Δ are equivalent, by Theorem 5.4.4, $\mathcal{B}^c$ and hence $\mathcal{B}$ are stabilizable if and only if all the complex zeros $\mu \in \mathbb{C}$ of $\mathcal{M}_{\gamma_n}$ are such that $|\mu| < 1$.

We first show that this condition is equivalent to the fact that all the quaternionic zeros $\lambda \in \mathbb{H}$ of $\mathcal{M}_{\gamma_n}$ satisfy $|\lambda| < 1$. Indeed, assume that the condition on the complex zeros of $\mathcal{M}_{\gamma_n}$ holds and let $\lambda \in \mathbb{H}\backslash\mathbb{C}$ be such that $\mathcal{M}_{\gamma_n}(\lambda) = 0$. By Lemma 2.2.10-1 and the definition (2.1) of $\psi_{[\lambda]}$, it follows that there exists $\mu \in [\lambda] \cap \mathbb{C}$. Since $\mathcal{M}_{\gamma_n} \in \mathbb{R}[s]$, if $\mathcal{M}_{\gamma_n}(\lambda) = 0$, also $\mathcal{M}_{\gamma_n}(\overline{\lambda}) = 0$ and, as $\lambda \neq \overline{\lambda} \sim \lambda$, by Lemma 2.2.10-3 it follows that $\mathcal{M}_{\gamma_n}(\mu) = 0$ too. Since by assumption $|\mu| < 1$ and $|\lambda| = |\mu|$, because $\mu \in [\lambda]$, we conclude that $|\lambda| < 1$. The reciprocal implication is obvious.

Now it just remains to prove that

$$\big(\mathcal{M}_{\gamma_n}(\nu) = 0 \text{ with } \nu \in \mathbb{H} \Rightarrow |\nu| < 1\big) \;\Leftrightarrow\; \big(\gamma_n(\lambda) = 0 \text{ with } \lambda \in \mathbb{H} \Rightarrow |\lambda| < 1\big).$$

"$\Rightarrow$" This implication is obvious since γ_n is, by definition, a right divisor $\mathcal{M}_{\gamma_n}$. Just note that if $\gamma_n(\lambda) = 0$, $\lambda \in \mathbb{H}$, then by Proposition 2.2.3 $\mathcal{M}_{\gamma_n}(\lambda) = 0$ and therefore $|\lambda| < 1$.

"$\Leftarrow$" Recall that by Lemma 2.2.21 he have

$$\overline{\gamma}_n \gamma_n = \mathcal{F}_{\gamma_n} \mathcal{M}_{\gamma_n}. \tag{5.13}$$

Let $\nu \in \mathbb{H}$ be such that $\mathcal{M}_{\gamma_n}(\nu) = 0$. This implies that $\mathcal{F}_{\gamma_n}\mathcal{M}_{\gamma_n}(\nu) = 0$ and by (5.13) we have that $\overline{\gamma}_n\gamma_n(\nu) = 0$. If $\gamma_n(\nu) = 0$ then $|\nu| < 1$ and the result follows. Otherwise, by Proposition 2.2.5 there exists $\lambda \sim \nu$ such that $\overline{\gamma}_n(\lambda) = 0$ and by Corollary 2.2.8 there exists $\lambda' \sim \lambda$ such that $\gamma_n(\lambda') = 0$. Since $|\nu| = |\lambda| = |\lambda'| < 1$ the statement is proved. $\square$

5.5 BIBO-stability

In the analysis of i/o systems, the most widely used concept is called *BIBO (bounded input-bounded output) stability.*

Definition 5.5.1. [43] An i/o behavior (4.8) is *BIBO-stable* if it does not contain trajectories with bounded input and unbounded output, i.e.,

$$\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathcal{B} \text{ and } ||u||_\infty < \infty \Rightarrow ||y||_\infty < \infty,$$

where $||u||_\infty = \sup\{||u(t)|| : t \in \mathbb{Z},\, t > 0\}$.

Remark 5.5.2. A state-space model is BIBO-stable in *classical* systems theory if bounded inputs generate bounded outputs *when the initial state is zero*. Clearly, if such a model is BIBO-stable in the *behavioral sense*, it is BIBO-stable in the *state-space sense*. The reciprocal fact is not true as we show in the next example. $\square$

Example 5.5.3. Consider the discrete-time i/o system

$$(\sigma - 2)y = (\sigma - 2)u. \tag{5.14}$$

The realization of the system given by

$$\begin{cases} \sigma x &= 2x \\ y &= x + u \end{cases}$$

easily shows that $x(t) = 2^t x(0)$ and therefore, if $x(0) = 0$, $y = u$. In the *classical sense* the system is BIBO-stable.

However, let $\mathcal{B}_{\mathrm{i/o}}$ be the i/o behavior of (5.14). If $u = 0$ and $y(t) = 2^t$ then $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathcal{B}_{\mathrm{i/o}}$, which is not BIBO-stable from a behavioral point of view since u is bounded and y is unbounded. Note that $\ker(\sigma - 2)$ is not stable. □

To generalize the situation evidenced by the latter example, consider an i/o quaternionic system with representation (4.8), and define

$$\tilde{\mathcal{B}}_{\mathrm{i/o}} = \left\{ \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \left(\mathbb{H}^{p+m}\right)^{\mathbb{Z}} \,:\, \tilde{P}y = \tilde{Q}u \right\}, \tag{5.15}$$

with $\tilde{P}$ and $\tilde{Q}$ such that $(\tilde{P}, \tilde{Q})$ are left coprime, $P = L\tilde{P}$ has full rank and $Q = L\tilde{Q}$ for some suitable matrix L.

Lemma 5.5.4. *The behavior $\mathcal{B}_{i/o}$ is BIBO-stable if and only if $\tilde{\mathcal{B}}_{i/o}$ is BIBO-stable and* $\ker L$ *is stable.*

Proof. "Only if". Assume that $\mathcal{B}_{\mathrm{i/o}}$ is BIBO-stable. Since $\tilde{\mathcal{B}}_{\mathrm{i/o}} \subseteq \mathcal{B}_{\mathrm{i/o}}$, also $\tilde{\mathcal{B}}_{\mathrm{i/o}}$ is BIBO-stable. If by contradiction $\ker L$ is unstable, there exists z unbounded such that $Lz = 0$. Since the operator $\tilde{P}$ is surjective, there exists y, necessarily unbounded, such that $z = \tilde{P}y$ and $\left[\begin{smallmatrix} y \\ 0 \end{smallmatrix}\right] \in \mathcal{B}_{\mathrm{i/o}}$, because $Py = L\tilde{P}y = Lz = 0$. Thus $\ker L$ must be stable.

"If". Assume that $\tilde{\mathcal{B}}_{\mathrm{i/o}}$ is BIBO-stable and $\ker L$ is stable. If $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathcal{B}_{\mathrm{i/o}}$, then $v = \tilde{P}y - \tilde{Q}u \in \ker L$ which is stable. Thus, $\tilde{P}y = \tilde{Q}u + v$ for some v bounded. As the matrices $\tilde{P}$ and $\tilde{Q}$ are left coprime, analogous to the commutative case, they satisfy a Bézout equation, i.e., there exist polynomial matrices S and T such that $\tilde{P}S = \tilde{Q}T + I$.

Applying these operators to v we obtain $\tilde{P}Sv = \tilde{Q}Tv + v$ and subtracting this equation from $\tilde{P}y = \tilde{Q}u + v$, we get $\tilde{P}(y - Sv) = \tilde{Q}(u - Tv)$. If u is bounded, so is $u - Tv$ and, by BIBO-stability of $\tilde{\mathcal{B}}_{\mathrm{i/o}}$, also $y - Sv$ and, consequently, y. □

The isomorphism between a behavior and its complex form leads to the following result.

Lemma 5.5.5. *An i/o behavior $\mathcal{B}$ is BIBO-stable if and only if its complex form, $\mathcal{B}^c$, is BIBO-stable.*

In the commutative case, BIBO-stability of an i/o behavior can be characterized in terms of the Smith-McMillan form of the transference matrix of the behavior. We present next a theorem which is a consequence of [43, Theorem 7.6.2].

Theorem 5.5.6. *Let $\mathcal{B}_{i/o} \in (\mathbb{C}^{p+m})^{\mathbb{Z}}$ be an i/o behavior with kernel representation $\begin{bmatrix} P & -Q \end{bmatrix}$ with P and Q left prime and let*

$$\operatorname{diag}\left(\frac{\epsilon_1}{\psi_1}, \ldots, \frac{\epsilon_p}{\psi_p}\right)$$

be the Smith-McMillan form of the transfer matrix $P^{-1}Q$. Then $\mathcal{B}_{i/o}$ is BIBO-stable if and only if ψ_1 is asymptotically stable.

Using the results so far obtained, we can now characterize BIBO-stable quaternionic systems.

Theorem 5.5.7. *Let $\mathcal{B}_{i/o}$ be the quaternionic i/o behavior* (4.8) *and*

$$\operatorname{diag}(\gamma_1, \ldots, \gamma_p) \ \textit{and} \ \operatorname{diag}\left(\frac{\epsilon_1}{\psi_1}, \ldots, \frac{\epsilon_p}{\psi_p}\right)$$

be a quaternionic Smith form of $\begin{bmatrix} P & -Q \end{bmatrix}$ and a quaternionic Smith-McMillan form of $P^{-1}Q$, respectively. Then $\mathcal{B}_{i/o}$ is BIBO-stable if and only if γ_p is stable and ψ_1 is asymptotically stable.

Proof. Define $\tilde{\mathcal{B}}_{\mathrm{i/o}}$, $\tilde{P}$, $\tilde{Q}$, and L as in (5.15) and let the behavior $\tilde{\mathcal{B}}^{\mathbb{C}}_{\mathrm{i/o}}$ be the complex form of $\tilde{\mathcal{B}}_{\mathrm{i/o}}$. Notice that, as mentioned in Section 4.2.3, $\tilde{\mathcal{B}}^{\mathbb{C}}_{\mathrm{i/o}}$ has kernel representation $\begin{bmatrix} \tilde{P}^c & -\tilde{Q}^c \end{bmatrix}$, i.e., transfer matrix $(\tilde{P}^c)^{-1}\tilde{Q}^c = (\tilde{P}^{-1}\tilde{Q})^c = (P^{-1}Q)^c$. By

Lemma 5.5.5 $\tilde{\mathcal{B}}_{\text{i/o}}$ is BIBO-stable if and only if $\tilde{\mathcal{B}}^{\mathbb{C}}_{\text{i/o}}$ is BIBO-stable and, by Theorems 5.5.6 and 3.5.9, $\tilde{\mathcal{B}}^{\mathbb{C}}_{\text{i/o}}$ is BIBO-stable if and only if $\mathcal{M}_{\psi_1}$ is asymptotically stable. Moreover, by Lemma 5.3.11 $\mathcal{M}_{\psi_1}$ is asymptotically stable if and only if ψ_1 is asymptotically stable. Therefore, by Lemma 5.5.4, the statement is proved if we show that $\ker L$ is stable if and only if γ_p is stable.

Since $\tilde{P}$ and $\tilde{Q}$ are left coprime, by Theorem 5.1.5, the quaternionic Smith form of $\begin{bmatrix}\tilde{P} & -\tilde{Q}\end{bmatrix}$ is $\begin{bmatrix}I & 0\end{bmatrix}$. So, if $\begin{bmatrix}S & 0\end{bmatrix}$ is a quaternionic Smith form of $\begin{bmatrix}P & -Q\end{bmatrix}$, there are unimodular matrices U, V, $\tilde{U}$, and $\tilde{V}$ such that

$$\begin{bmatrix}S & 0\end{bmatrix} = U\begin{bmatrix}P & -Q\end{bmatrix}V = UL\begin{bmatrix}\tilde{P} & -\tilde{Q}\end{bmatrix}V = UL\tilde{U}^{-1}\begin{bmatrix}I & 0\end{bmatrix}\tilde{V}^{-1}V$$

and therefore, if $\begin{bmatrix}X & Y\end{bmatrix}$ are the first p rows of $\tilde{V}^{-1}V$,

$$\begin{bmatrix}S & 0\end{bmatrix} = UL\tilde{U}^{-1}\begin{bmatrix}X & Y\end{bmatrix}. \tag{5.16}$$

From this we obtain that $0 = UL\tilde{U}^{-1}Y$. However, L has full rank, therefore $Y = 0$. It follows that the unimodular matrix $\tilde{V}^{-1}V$ is block triangular, hence X (which is a block on the diagonal of $\tilde{V}^{-1}V$) must be unimodular. Now, from (5.16), $S = UL\tilde{U}^{-1}X$ is a quaternionic Smith form of L and this, by Theorem 5.3.12, concludes the proof. □

Remark 5.5.8. By its definition, BIBO-stability in the *classical sense* only considers the i/o relation (i.e., the transfer matrix $P^{-1}Q$), thus ignoring the *internal behavior* of the state. Therefore, in the notation of Theorem 5.5.7, this property is equivalent to asymptotic stability of ψ_1. □

5.6 Observability

Observability expresses the possibility of obtaining information concerning some components of a trajectory by observing the values of the other ones.

Definition 5.6.1. [43] Let $\Sigma = (\mathbb{T}, \mathbb{W}, \mathcal{B})$ be a time-invariant dynamical system and suppose that the trajectories in $\mathcal{B}$ are partitioned as $w = (w_1, w_2)$. We say that w_2 is *observable* from w_1 if $(w_1, w_2), (w_1, w_2') \in \mathcal{B}$ implies that $w_2 = w_2'$.

Clearly, for linear behaviors $\mathcal{B}$, w_2 is observable from w_1 if and only if $(0, w_2) \in \mathcal{B}$ implies that $w_2 = 0$. In particular, if $\mathcal{B}$ is given as $R_1(\sigma, \sigma^{-1})w_1 = R_2(\sigma, \sigma^{-1})w_2$, then w_2 is observable from w_1 if and only if $\ker R_2(\sigma, \sigma^{-1}) = \{0\}$.

The next result relates the observability in a quaternionic system with the observability in its complex form.

Lemma 5.6.2. *Let $R_1 \in \mathbb{H}^{g\times r_1}[s, s^{-1}]$ and $R_2 \in \mathbb{H}^{g\times r_2}[s, s^{-1}]$. Let $\mathcal{B}$ be the behavior defined by $R_1(\sigma, \sigma^{-1})w_1 = R_2(\sigma, \sigma^{-1})w_2$. Then the variable w_2 is observable from w_1 if and only if $w_2^{\mathbb{C}}$ is observable from $w_1^{\mathbb{C}}$.*

Proof. This result follows immediately from the definitions of observability and of the complex form $\mathcal{B}^{\mathbb{C}}$ of $\mathcal{B}$. □

Analogously to the commutative case observability is characterized as follows.

Theorem 5.6.3. *Let $R_1 \in \mathbb{H}^{g\times r_1}[s, s^{-1}]$ and $R_2 \in \mathbb{H}^{g\times r_2}[s, s^{-1}]$. Let $\mathcal{B}$ be the behavior defined by $R_1(\sigma, \sigma^{-1})w_1 = R_2(\sigma, \sigma^{-1})w_2$. Then the following conditions are equivalent:*

(i) w_2 is observable from w_1;

(ii) R_2 is right prime;

(iii) the quaternionic Smith form of R_2 is $\begin{bmatrix} I_{r_2} \\ 0 \end{bmatrix}$.

Proof. We will show that the implications $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$ hold true.

$(i) \Rightarrow (ii)$ Suppose that R_2 is not right prime, i.e., there exists a non unimodular matrix D such that $R_2 = \widetilde{R}_2 D$. Since D is not unimodular there exists a nonzero trajectory w_2 such that $D(\sigma, \sigma^{-1})w_2 = 0$. Thus,

$$R_2(\sigma, \sigma^{-1})w_2 = \widetilde{R}_2(\sigma, \sigma^{-1})D(\sigma, \sigma^{-1})w_2 = 0,$$

which means that $\ker R_2 \neq \{0\}$ and hence w_2 is not observable from w_1.

$(ii) \Rightarrow (iii)$ Let R_2 be right prime. Then R_2^T is left prime, and by Theorem 5.1.5 we have that the quaternionic Smith form of R_2^T is $\begin{bmatrix} I_{r_2} & 0 \end{bmatrix}$. Therefore, the quaternionic Smith form of R_2 is $\left[\begin{smallmatrix} I_{r_2} \\ 0 \end{smallmatrix}\right]$.

$(iii) \Rightarrow (i)$ Suppose that the quaternionic Smith form of R_2 is $\left[\begin{smallmatrix} I_{r_2} \\ 0 \end{smallmatrix}\right]$. Then there exist unimodular matrices $U \in \mathbb{H}^{g \times g}[s, s^{-1}]$ and $V \in \mathbb{H}^{r_2 \times r_2}[s, s^{-1}]$ such that $UR_2V = \left[\begin{smallmatrix} I_{r_2} \\ 0 \end{smallmatrix}\right]$. Let $w_2 \in \ker R_2$, i.e., $R_2(\sigma, \sigma^{-1})w_2 = 0$. Then, $U(\sigma, \sigma^{-1})R_2(\sigma, \sigma^{-1})w_2 = 0$, or equivalently,

$$\begin{bmatrix} I_{r_2} \\ 0 \end{bmatrix} V^{-1}(\sigma, \sigma^{-1})w_2 = 0 \qquad \text{since} \qquad UR_2 = \begin{bmatrix} I_{r_2} \\ 0 \end{bmatrix} V^{-1}.$$

But this implies that $V^{-1}(\sigma, \sigma^{-1})w_2 = 0$, and because V is unimodular, we have that $w_2 = 0$, showing that $\ker R_2(\sigma, \sigma^{-1}) = \{0\}$. $\square$

Conclusions

The behavioral approach to dynamical systems, introduced by J. C. Willems [53, 54] in the eighties, considers as the main object of study in a system the set of all the trajectories which are compatible with its laws, known as the system behavior. In Chapter 4 we extended this approach to quaternionic systems. In particular, we studied behaviors that can be described as solution sets of quaternionic matrix difference equations, i.e., those which are the kernel of some suitable matrix difference operator. To study properly this class of systems we had first to give some background material on quaternions and quaternionic polynomials.

After have given, in Chapter 1, some notions on quaternions and quaternionic matrices we dedicated the second chapter to the study of quaternionic polynomials. The nontrivial relation between the zeros and factors of such polynomials was pointed out. Moreover, new definitions of total divisor and similarity of quaternionic polynomials were presented and their relation to other existing definitions was studied. These notions were relevant in Chapter 3 where quaternionic polynomial and rational matrices have been studied. In that chapter a new definition of determinant for quaternionic polynomial matrices, Pdet, was proposed. Given a square matrix R, the polynomial determinant $\mathrm{Pdet}(R)$ was defined as

$$\mathrm{Pdet}(R) = \prod_{l=1}^{n} \gamma_l \overline{\gamma}_l,$$

where the γ_l's are the main diagonal elements of any triangular matrix obtained from R pre-multiplying it by a matrix $U \in SL(n, \mathbb{H}[s])$. This polynomial determinant was used in the last chapter to characterize the stability of quaternionic behavioral systems. Also in Chapter 3, complex adjoint matrices were defined and shown to share many algebraic properties with the corresponding quaternionic polynomial matrices. Furthermore, we introduced the quaternionic Smith form of quaternionic polynomial matrices; we also characterized the complex Smith form of complex adjoint matrices and gave its relation with the quaternionic Smith form of the corresponding quaternionic matrices. Finally, the quaternionic Smith-McMillan form of a quaternionic rational matrix was defined and the results on the Smith form were extended to rational matrices. These algebraic tools played an important role in Chapter 5 where we have characterized fundamental dynamical properties of quaternionic behaviors such as controllability and stability. Most of the results of the last chapter are analogous to the real or complex case. The main differences occur when we have to evaluate polynomials or polynomial matrices since evaluation of quaternionic polynomials is not a ring homomorphism.

As we have seen in Section 5.1, the well-known Hautus criterion to check the controllability of a system is not valid in the quaternionic case. In the future work we intend to investigate a quaternionic version of the Hautus criterion. Another open question is the characterization of (non asymptotic) stability in terms of the polynomial determinant Pdet, taking into account the geometric and algebraic multiplicity of its zeros.

References

[1] H. Aslaksen. Quaternionic determinants. *Math. Intelligencer*, 18(3):57–65, 1996.

[2] A. Bagazgoitia. A determinantal identity for quaternions. In *Linear algebra and applications (Vitoria-Gasteiz, 1983)*, pages 127–132. Univ. País Vasco-Euskal Herriko Unib., Bilbao, 1984.

[3] B. Beck. Sur les équations polynomiales dans les quaternions. *Enseign. Math. (2)*, 25(3-4):193–201, 1979.

[4] J. L. Brenner. Matrices of quaternions. *Pacific J. Math.*, 1:329–335, 1951.

[5] A. Cayley. On certain results relating to quaternions. *Philosophical Magazine*, 26:141–145, 1845.

[6] L. Cerlienco and M. Mureddu. A note on polynomial equations in quaternions. *Rend. Sem. Fac. Sci. Univ. Cagliari*, 51(1):95–99, 1981.

[7] N. Cohen and S. De Leo. The quaternionic determinant. *Electron. J. Linear Algebra*, 7:100–111 (electronic), 2000.

[8] P. M. Cohn. *Skew fields*, volume 57 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995.

[9] S. De Leo and G. C. Ducati. Real linear quaternionic differential operators. *Comput. Math. Appl.*, 48(12):1893–1903, 2004.

[10] J. E. Dennis, Jr., J. F. Traub, and R. P. Weber. The algebraic theory of matrix polynomials. *SIAM J. Numer. Anal.*, 13(6):831–845, 1976.

[11] J. Dieudonné. Les déterminants sur un corps non commutatif. *Bull. Soc. Math. Fr.*, 71:27–45, 1943.

[12] A. El-Gohary and E. R. Elazab. Exponential control of a rotational motion of a rigid body using quaternions. *Appl. Math. Comput.*, 137(2-3):195–207, 2003.

[13] E. Fornasini. *Dispensa Di Sistemi Multivariabili.* University of Padua, 2002.

[14] F. G. Frobenius. Über lineare substitutionen und bilineare formen. *Für die reine und angewandte Mathematik*, 84:1–63, 1878.

[15] F. R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998.

[16] B. Gordon and T. S. Motzkin. On the zeros of polynomials over division rings. *Trans. Amer. Math. Soc.*, 116:218–226, 1965.

[17] R. M. Guralnick, L. S. Levy, and C. Odenthal. Elementary divisor theorem for noncommutative PIDs. *Proc. Amer. Math. Soc.*, 103(4):1003–1011, 1988.

[18] S. W. R. Hamilton. Theory of conjugate functions, or algebraic couples; with a preliminary and elementary essay on algebra as the science of pure time. *Transactions of the Royal Irish Academy*, 17:293–422, 1837.

[19] S. W. R. Hamilton. *Elements of Quaternions. Second Edition.* Longman, 1889.

[20] M. Hazewinkel, J. Lewis, and C. Martin. Symmetric systems with semisimple structure algebra: the quaternionic case. *Systems Control Lett.*, 3(3):151–154, 1983.

[21] D. Hestenes. *New foundations for classical mechanics.* Kluwer Academic Publishers Group, Dordrecht, The Netherlands, 1986.

[22] B. Jacob. *Linear Algebra.* W. H. Freeman and Company, New York, 1990.

[23] N. Jacobson. *The Theory of Rings.* American Mathematical Society, New York, 1943.

[24] N. Jacobson. *Basic algebra. I.* W. H. Freeman and Company, New York, 1985.

[25] T. Kailath. *Linear Systems.* Prentice Hall, Englewood Cliffs, 1980.

[26] J. B. Kuipers. *Quaternions and rotation sequences.* Princeton University Press, Princeton, NJ, 1999.

[27] T. Y. Lam. *A first course in noncommutative rings.* Springer-Verlag, New York, 1991.

[28] P. Lancaster and M. Tismenetsky. *The theory of matrices.* Computer Science and Applied Mathematics. Academic Press Inc., Orlando, FL, 1985.

[29] S. Lang. *Algebra.* Addison-Wesley, Reading, Massachusetts, 3rd edition, 1993.

[30] H. C. Lee. Eigenvalues and canonical forms of matrices with quaternion coefficients. *Proc. Roy. Irish Acad. Sect. A.*, 52:253–260, 1949.

[31] S. D. Leo, G. C. Ducati, and C. C. Nishi. Quaternionic potentials in non-relativistic quantum mechanics. *Journal of Physics A: Mathematical and General*, 35(26):5411–5426, 2002.

[32] L. S. Levy and J. C. Robson. Matrices and pairs of modules. *J. Algebra*, 29:427–454, 1974.

[33] R. Mukundan. Quaternions: From classical mechanics to computer graphics, and beyond. In *Proocedings of the* 7^{th} *Asian Technology Conference in Mathematics*, 2002.

[34] T. Nakayama. A note on the elementary divisor theory in non-commutative domains. *Bull. Amer. Math. Soc.*, 44:719–723, 1938.

[35] I. Niven. Equations in quaternions. *Amer. Math. Monthly*, 48:654–661, 1941.

[36] I. Niven. The roots of a quaternion. *Amer. Math. Monthly*, 49:386–388, 1942.

[37] G. Oliveira. Corpos não comutativos. *Boletim da Sociedade Portuguesa da Matemática*, 47:49–62, 2002.

[38] O. Ore. Linear equations in non-commutative fields. *Ann. of Math. (2)*, 32(3):463–477, 1931.

[39] R. Pereira and P. Rocha. Quaternionic behavioral systems. Technical Report CM03/I-17, Departamento de Matemática, Universidade de Aveiro, Portugal, 2003.

[40] R. Pereira and P. Rocha. On the determinant of quaternionic polynomial matrices and its application to system stability. *Submitted for publication in Linear Algebra and its Applications*, 2005.

[41] R. Pereira, P. Rocha, and P. Vettori. Algebraic tools for the study of quaternionic behavioral systems. *Linear Algebra and its Applications*, 400:121–140, 2005.

[42] R. Pereira and P. Vettori. Stability of quaternionic linear systems. *To appear in IEEE Trans. Automat. Control*, 51(3), March 2006.

[43] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach*, volume 26 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 1997.

[44] S. Pumplün and S. Walcher. On the zeros of polynomials over quaternions. *Comm. Algebra*, 30(8):4007–4018, 2002.

[45] H. H. Rosenbrock. *State-space and multivariable theory.* John Wiley & Sons, Inc. [Wiley Interscience Division], New York, 1970.

[46] K. Shoemake. Animating rotation with quaternion curves. *Computer Graphics*, 19(3):245–254, 1985.

[47] E. Study. Zur theorie der linearen gleichungen. *Acta Math.*, 42:1–61, 1920.

[48] O. Teichmüller. *Der Elementarteilersatz Für Nichtkommutative Ringe.* S.-B.-Preuss Akad. Wiss., 1937.

[49] M. Vidyasagar. *Control system synthesis.* MIT Press, Cambridge, MA, 1985.

[50] P. B. Visscher and X. Feng. Quaternion-based algorithm for micromagnetics. *Phys. Rev. B*, 65(104412), 2002.

[51] J. Vrbik. A novel solution to Kepler's problem. *European J. Phys.*, 24(6):575–583, 2003.

[52] J. C. Willems. From time series to linear system — part I. finite dimensional linear time invariant systems. *Automatica—J. IFAC*, 22(5):561–580, 1986.

[53] J. C. Willems. Models for dynamics. In U. Kirchgraber and H. O. Walther, editors, *Dynamics Reported*, volume 2, pages 171–269. John Wiley & Sons Ltd., Chichester, 1989.

[54] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Automat. Control*, 36(3):259–294, Mar. 1991.

[55] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra Appl.*, 251:21–57, 1997.

[56] Z. Y. Zhang. An algebraic principle for the stability of difference operators. *J. Differential Equations*, 136(2):236–247, May 1997.